Univerzita Karlova v Praze
Přírodovědecká fakulta

Studijní program: Parazitologie / Study Programme: Parasitology

Mgr. Anna Dostálová (Svárovská)

Leishmania attachment in permissive vectors and the role of sand fly midgut proteins in parasite-vector interaction

Přichycení leishmanií ve střevě permisivních přenašečů a role střevních proteinů flebotomů v interakci parazit-přenašeč

PhD Thesis / Dizertační práce

Supervisor / Školitel: Prof. RNDr. Petr Volf, CSc.

Praha, 2011

Authors`s declaration / Prohlášení autorky:

I declare that the submitted thesis is my own work and that I properly cited all scientific literature used. Neither this thesis as a whole nor its substantial part has been submitted for the award of any other degree or diploma.

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

Anna Dostálová
26.4.2011, Praha

Supervisor`s declaration / Prohlášení školitele:

I declare that Anna Dostálová substantially contributed to the experimental work in the three projects presented in her thesis and had a principal role in writing two of the three publications presented.

Prohlašuji, že Anna Dostálová se významně podílela na experimentální práci na třech projektech shrnutých v této dizertační práci a je hlavní autorkou textu dvou publikací.

Prof. RNDr. Petr Volf, CSc.
26.4.2011, Praha

Acknowledgements

I wish to thank:

my supervisor Petr Volf for his encouragment and unceasing support throughout my PhD studies

members of the vector team for making the lab such a great place to work. Special thanks go to Lucka Ječná and Jitka Pecková, who participated in the projects presented in this thesis and to Iva Rohoušová for her incredible patience.

Helena Kulíková, for all the help that went well beyond her official task of administrative support

Ryan Jochim for his ‘contagious’ enthusiasm for science

Jesus Valenzuela, Paul Bates and their teams for being such excellent hosts during my internships in their labs

my husband, my whole family and God for always standing by my side.

Abstract

This thesis focuses on the development of protozoan parasites of the genus *Leishmania* in their insect vectors, sand flies. It sums up results of three projects I was involved in during my PhD studies. Main emphasis was put on permissive sand fly species that support development of various species of *Leishmania*. Using a novel method of binding of fluorescently labeled leishmania promastigotes to the midguts *in vitro*, we studied the stage- and species-specificity of the binding. We demonstrated that *Leishmania* midgut binding is strictly stage-dependent, is a property of forms found in the middle phase of development (long and short nectomonad forms), but is absent in early forms occurring in within the blood meal, procyclics, and in final stages, metacyclics. Comparing the binding of several leishmania species, we showed the natural parasite is not necessarily the species that can always bind *in vitro* most efficiently to the midgut of its vector. In some cases, we even observed significant binding of *Leishmania* species that do not survive in the midgut of the particular sand fly species *in vivo*. We conclude that the specificity of *in vitro* binding alone is insufficient to explain overall vector specificity. In our second study we looked into the role of *Leishmania* surface phosphoglycans during parasite development in three sand fly species. We showed an essential role of *Leishmania major* surface lipophosphoglycan in its natural vector *P. duboscqi*, while we observed complete development of parasites devoid of lipophosphoglycan in two permissive vectors, *P. perniciosus* and *P. argentipes*. We also demonstrated that *P. duboscqi* is not able to support development of *L. infantum* and is therefore classified as a specific vector. Our results give important supporting evidence for the presence of a lipofosfoglycan-independent mechanism of attachment in permissive sand flies. In the third study, we focused on molecules expressed in the midgut of the permissive vector, *P. perniciosus*, which are most likely to be involved in interaction with transmitted parasites. We constructed, sequenced and analyzed two midgut specific cDNA libraries from sugar fed and blood fed female *P. perniciosus* and compared the transcript expression profiles. By detailed comparison of our findings with the published midgut transcriptomes of two other sand fly species we identified several features potentially important for their vectorial competence and shared by the two permissive vectors, *P. perniciosus* and *Lutzomyia longipalpis*.

Abstrakt

Tato práce se zabývá vývojem prvoků rodu *Leishmania* v jejich hmyzích přenašečích, flebotomech. Shrnuje výsledky tří projektů, na kterých jsem se během svého doktorského studia podílela. Zaměřila jsem se především na permisivní druhy flebotomů, tedy takové, které umožňují vývoj mnoha druhů lieshmanií. S využitím nové metody vazby fluorescenčně značených leishmanií na střevo flebotomů *in vitro* jsme sledovali vazbu různých morfologických stádií a druhů leishmanií. Nutnost přichytit se ke střevnímu epitelu je považována za podmínku úspěšného vývoje leishmanií v jejich přenašečích a důležitý faktor určující specifitu vztahu parazit-přenašeč. Zjistili jsme, že na střeva se vážou takřka výhradně tzv. dlouhé nektomonády a leptomonády, tedy stádia nacházená ve střední fázi infekce. Naopak procyklická stádia, tedy formy nacházené v přenašeči krátce po infekci, nebyla schopná se vázat. Podobné zjištení platí i pro metacyklické promastigoty vyskytující se v pozdní fázi. Srovnání vazby několika druhů leishmanií ukázalo, že ne vždy se parazit přenášený daným vektorem váže nejlépe *in vitro*. Překvapivě jsme pozorovali také vazbu některých druhů leishmanií i na střeva „nekompatibilních" druhů specifických přenašečů, ve kterých daná leishmanie není schopna vývoje. Zdá se tedy, že samotná schopnost vazby *in vitro* nevysvětluje specifitu některých druhů flebotomů k přenášeným druhům leishmanií. V další studii jsme zkoumali roli povrchových fosfoglykanů leishmanií při přichycení a vývoji ve třech druzích flebotomů. Prokázali jsme zásadní roli lipofosfoglykanu při přichycení *Leishmania major* v jejím specifickém přenašeči *Phlebotomus duboscqi* a naopak jeho postradatelnost při vývoji ve dvou permisivních druzích, *P. perniciosus* a *P. argentipes*. Popsali jsme také, že *P. duboscqi* nepodporuje vývoj *L. infantum* a charakterizovali ho tedy jako specifického přenašeče. Společně s výsledky předchozích studií tedy naše data svědčí o tom, že mechanismus přichycení leishmanií ve střevě permisivních vektorů je nezávislý na lipofosfoglykanu.. Ve třetí práci jsme popsali střevní transkriptom *P. perniciosus* a diskutujeme možné role popsaných molekul při interakcích s leishmaniemi. Zkonstruovali a analyzovali jsme dvě střevní cDNA knihovny a srovnávali molekuly exprimované ve střevě *P. perniciosus* před a po sání krve. Analyzovali jsme také vztahy k dříve popsaným střevním transkriptomům jiných dvou druhů flebotomů a identifikovali rozdíly a podobnosti, které mohou hrát roli v rozdílné specifitě těchto přenašečů vůči leishmaniím.

Table of contents

1 Introduction

1.1 *Leishmania* development in the sand fly midgut and transmission

Phlebotomine sand flies (Diptera: Psychodidae, subfamily Phlebotominae) are small (1.5 – 3mm body length) insects that are principally found in tropical and subtropical regions. Both males and females feed on natural sugar sources whereas only females feed on blood which provides nutrition for the production of eggs. Females of two sand fly genera, *Phlebotomus* and *Lutzomyia*, are of medical importance as the only proven vectors of *Leishmania* species pathogenic for humans (Killick-Kendrick 1999).

The genus *Leishmania* (Kinetoplatida: Trypanosomatidae) are protozoan parasites causing a spectrum of diseases called leishmaniases in their vertebrate hosts, including humans. There are about 10 *Leishmania* species of significant importance for public health. Symptoms of leishmaniases can range from mild self-healing cutaneous lesions to fatal visceral cases. The lack of a human vaccine, increasing resistance to the currently used drugs and their serious side effects urge the need for research of leishmaniasis.

The parasite has a digenetic life-cycle alternating between a mammalian host and a sand fly vector. In this thesis I will focus on the development in the sand fly. Studies elucidating parasite-vector interactions have become the basis for novel approaches to reduce transmission of several insect-borne diseases. They provide targets for use as novel vector-based transmission-blocking vaccines (TBVs) (Coutinho-Abreu and Ramalho-Ortigao 2010). TBVs aim at preventing the transmission of pathogens by targeting molecule(s) expressed on the surface of pathogens during their developmental phase within the insect vector or by targeting molecules expressed by the vectors. This strategy has been used successfully in identifying promising vaccine candidates for malaria control (Dinglasan and Jacobs-Lorena 2008) and have a great potential in the research of leishmaniases.

Outside the mammalian host, the *Leishmania* life cycle is confined to the digestive tract of sand flies. The precise location differs between subgenera *Leishmania* and *Viannia*. The New World subgenus *Viannia*, e.g. *Leishmania braziliensis*, enter the hindgut before migrating forward into the midgut and are therefore called peripylarian

parasites (Lainson et al. 1977). However, most *Leishmania* species (subgenus *Leishmania*) are suprapylarian parasites; that is, their development is restricted to the midgut (reviewed by Sadlova 1999). As most studies on parasite–vector interactions and also my experimental work have focused on suprapylarian parasites, below I discuss their development in more detail.

Infection is initiated when female sand flies ingest blood that contains macrophages infected with amastigotes, small (3–5 μm), an immotile and rounded form of the parasite. The change in conditions moving from the mammalian host to the sand fly midgut (such as decrease in temperature and increase in pH) triggers development of the parasite in the vector. The amastigotes transform into procyclic promastigotes - weakly motile forms with a short flagellum beating at the anterior end of the cell. These are the first replicative forms that proliferate in the early bloodmeal and are separated from the midgut by a type I peritrophic matrix (reviewed by Kamhawi 2006). Around 48–72 hours later, parasites begin to slow their replication and differentiate into non-dividing, strongly motile long nectomonad promastigotes. These escape from the peritrophic matrix-encased blood meal into the midgut lumen. They move towards the anterior midgut and later develop into short nectomonad promastigotes (called leptomonads by Rogers et al. 2002), which enter another proliferative cycle (Gossage et al. 2003, Kamhawi 2006, Bates 2007).

Ultimately, *Leishmania* transform into infective metacyclic stages and their delivery to the skin of the vertabrate host must be ensured for effective transmission. Metacyclics are small, rapid-swimming forms with an elongated flagellum that originate from leptomonads. It has been shown that metacyclogenesis in *Leishmania* is induced *in vitro* by low pH and nutrient depletion, while reduced tetrahydrobiopterin levels may also act as a signal for parasite differentiation (Cunningham et al. 2001, Bates 2008). However, very little is known about the actual signals triggering metacyclogenesis in the sand fly midgut. Of interest, a V-ATPase has been recently described in the midgut of *L. longipalpis* that could be involved in gut acidification (Ramalho-Ortigao et al. 2007a).

Occasionally, *Leishmania* metacyclics were observed in salivary gland of sand flies (Killick-Kendrick et al. 1996) or in urine droplets discharged by infected females during blood feeding (Sadlova and Volf 1999). However, it is generally accepted that there are two main mechanism of transmission of metacyclic parasites: either a limited number of metacyclics occuring in the proboscis is deposited into the skin during

feeding (Adler and Theodor 1935) or parasites residing behind the stomodeal valve (the junction between anterior midgut and foregut) are regurgitated with a backflow of ingested blood. Originally, the regurgitation was supposed to result from the mechanical block of the foregut or the stomodeal valve (Warburg and Schlein 1986). More recently, the damage to the chitin layer of the stomodeal valve (Schlein et al. 1991) and the role of parasite proteophosphoglycan were described. During late-stage infections, large numbers of short nectomonad and metacyclic parasites accumulate in the anterior midgut.. These parasites are packed in philamentous proteophosphoglycan (secreted most probably by the short nectomonad stages) creating a gel-like plug physically obstructing the gut (Rogers et al. 2004, Rogers and Bates 2007). Further contributing to the blockage of the gut is another form of the parasite called haptomonad forms. These leaf-like parasites are attached to cuticular lining of the stomodeal valve through an expanded flagellar tip containing hemidesmosomal structures. The ultrastructure of these hemidesmosomes has been known for many years, but it remains to be biochemically described both at the parasite and the vector side. The attached parasites cause damage to the structure of the stomodeal valve, likely interfering with its function and facilitating reflux of parasites from the midgut (Schlein et al. 1992, Volf et al. 2004). The destruction is likely due to the action of parasite secreted chitinase (Rogers et al. 2008). In a recent study, (Kimblin et al. 2008) performed quantification of *L. major* promastigotes deposited to the skin by single *Phlebotomus duboscqi* females. They observed a bimodal distribution of the numbers of transmitted parasites. Most of the sand flies delivered a low infectious dose (<600 parasites), while the remainder transmitted much higher doses, corresponding also to a higher percentage of the parasites present in the midgut before bloodfeeding. The authors suggest that this bimodality reflects the two distinct mechanisms of transmission (Kimblin et al. 2008).

1.2 Sand fly midgut proteins

Several natural barriers to *Leishmania* development have been described within the sand fly midgut. These include secreted proteolytic enzymes, the peritrophic matrix surrounding the ingested blood meal, sand fly immune reaction and the necessity to bind to the midgut epithelium. Surprisingly little is known about the molecular background of the parasite-vector interactions. While the genome sequences of several *Leishmania* species have been published (Peacock et al. 2007) and molecular studies

abound, molecular data on sand flies are limited. An analysis of expressed sequence tags (ESTs) from the whole *Lutzomyia longipalpis* sand fly and salivary gland transcriptomes of several sand fly species have been published (Dillon et al. 2006, Oliveira et al. 2009). With regard to *Leishmania* development in the midgut, particularly midgut-specific transcriptomic analyses of *L. longipalpis* and *Phlebotomus papatasi* (Ramalho-Ortigao et al. 2007b, Jochim et al. 2008, Pitaluga et al. 2009) have brought important insights into the repertoire of molecules expressed in the midgut. In the following paragraphs, I give an overview of sand fly midgut proteins that have been implicated in the parasite-vector interactions, with the exception of molecules involved in parasite attachment to the midgut epithelium that are discussed in a separate chapter (1.3).

1.2.1 Proteolytic enzymes

In the midgut of unfed sand flies there is little baseline protease activity. It is the ingestion of bloodmeal that induces secretion of digestive enzymes. Significant levels of protease activity are detected 6h PBM and peak levels are reached 18-48h PBM depending on the sand fly species. The pH optimum of the general endoprotease activity is in the alkaline range (7.5-9.5). Based on the results of assays using specific inhibitors, it can be attributed to serine proteases, namely trypsin- and chymotrypsin-like enzymes (Dillon and Lane 1993a, Telleria et al. 2010).

More recently, sequencing of ESTs has identified numerous transcripts coding for putative digestive enzymes in *P. papatasi* and *L. longipalpis*, the most abundant of them indeed being trypsins and chymotrypsins (Dillon et al. 2006, Ramalho-Ortigao et al. 2007b, Jochim et al. 2008, Pitaluga et al. 2009). Four putative trypsins have been described to date in the midgut of *L. longipalpis* and four in *P. papatasi*. They are all predicted as secreted proteins with molecular weight of the mature protein ranging from 25.9 to 27.9 kDa. The temporal expression profiles of the transcripts vary, as revealed by semi-quantitative end-point PCR (Ramalho-Ortigao et al. 2003, Telleria et al. 2007) and/or comparison of transcript abundance in cDNA libraries before and after blood feeding (Ramalho-Ortigao et al. 2007b, Jochim et al. 2008). In both species, one or several trypsin transcripts are present in high abundance in sugar fed females while their quantities decrease after the intake of blood. At the same time, the expression of other putative trypsins is induced upon blood feeding. Recently, Telleria et al. (2010) have

detected the expression of one of the *L. longipalpis* midgut trypsins (LlTryp1) also at the protein level. The observed protein expression profile corresponded to the transcript levels detected previously, confirming LlTryp1 as a blood-feeding induced molecule.

The second most abundant digestive enyzmes in the sand fly midgut are chymotrypsins. Three putative chymotrypsins have been identified in *P. papatasi* and six in *L. longipalpis* midgut. The expression patterns of these chymotrypsins are similar to the above mentioned trypsin molecules indicating that there may be early and late classes of serine proteases in sand flies, similar to what has been observed in some mosquitoes (Graf and Briegel 1989, Barillas-Mury et al. 1995). Trypsin and chymotrypsin transcripts abundant in the sugar fed sand fly midgut may be quickly translated and processed after blood feeding accounting for the initial blood meal digestion. Alternatively, inactive protein precursors can be stored in the midgut and processed after the intake of blood.

Besides trypsins and chymotrypsins, transcripts for other putative digestive proteases have also been described in the sand fly midgut. These include metallocarboxypeptidases, astacin-like metalloproteases and novel serine protease (Oliveira et al. 2009). Aminopeptidase activity has also been detected in the sand fly midgut after blood feeding, mainly associated with the midgut epithelium (Dillon and Lane 1993a).

It has been recognized for a long time that the activity of digestive enzymes affects *Leishmania* development in sand flies. Adler (1938) was the first to investigate how bloodmeal digestion during the early stages of infection in *P. papatasi*, which normally transmits *L. major*, might explain its resistance to other *Leishmania* species. He found that by decreasing the percentage of serum in the bloodmeal, the infection rate in flies infected with an inappropriate species (presumably *Leishmania tropica*) was significantly enhanced. He suggested that products of serum digestion destroy *Leishmania* parasites and that lowering the concentration of the serum protein components of the bloodmeal lowered the level of induced proteases. Several later studies showed reduced parasite numbers and even dead or destroyed parasites in the midguts of 'noncompatible' sand fly species in the early phase of infections, that is the time of the onslaught of the proteolytic activity (Shatova et al. 1984, Lawyer et al. 1990, Schlein and Jacobson 1998). Similarly to the initial Adler`s experiments, one of these studies (Schlein and Jacobson 1998) also reported enhanced survival of *L. donovani* in *P. papatasi* following meals devoid of serum and showed that this was correlated with

delayed timing and decreased levels of peak protease activities. Moreover, other studies revealed that even in 'compatible' parasite-vector combinations, up to 50% of the initial amastigote parasite inoculum is killed within the first day after blood feeding (Pimenta et al. 1997, Rogers et al. 2002).

Several publications have identified the digestive enzymes as one of the culprits of these early parasite losses using various methods to supress midgut proteolytic activity. The addition of soybean trypsin inhibitor to the bloodmeal promoted the early survivor of *L. donovani* in *P. papatasi* (Borovsky and Schlein 1987). The addition of the inhibitor also enhances survival of "compatible" parasites in sand flies, in which the formation of the peritrophic matrix has been blocked by chitinase (Pimenta et al. 1997, Rogers et al. 2008). The peritrophic matrix is under normal conditions thought to act as a partial barrier limiting the exposure of the parasites to the digestive enyzmes in the earliest phases of infection (for more details see section 1.2.2). Volf et al. (2001) report an enhancing effect of heparin on *L. major* infections in its natural vector *P. duboscqi*. The authors attribute this effect to the fact that heparin supressed midgut trypsin activity, while it did not affect defecation, oviposition or mortality of the sand flies. Recently, (Sant'Anna et al. 2009) have demonstrated that specific knock-down of a blood feeding induced trypsin (LlTryp1) in *L. longipalpis* promotes the survival of *L. mexicana*.

In order to complete its developmental cycle, *Leishmania* must have developed mechanisms to overcome the hostile environment of the blood fed midgut. Numerous studies provide evidence the parasite manipulates the levels and timing of protease activity in the midgut. Schlein and Romano (1986) and Dillon and Lane (1993b) demonstrated the ability of *L. major* to suppress or delay the peak of trypsin and aminopeptidase activity in the midgut. Similar observations have recently been made in *L. longipalpis*. Both *L. mexicana* (Sant'Anna et al. 2009) and *L. infantum* (Telleria et al. 2010) infections led to a decreased trypsin activity in the midgut. However, it should be noted that the survival of *L. major* in *P. papatasi* and *P. duboscqi* has also been observed in the absence of any significant inhibition of or delay in peak protease activities during infection (Pimenta et al. 1997, Secundino et al. 2010). In each of the above cited studies, one synthetic substrate was used to measure the overall trypsin activity (benzoyl-ariginine-p-nitroanilid in most cases) and one for chymotrypsin activity in some.

Recent transcriptomic studies have revealed that the presence of *L. major* and *L. infantum* in the midgut of *P. papatasi* and *L. longipalpis*, respectively, can affect the abundance of several digestive enzymes transcripts after blood feeding (Ramalho-Ortigao et al. 2007b, Jochim et al. 2008). These observations suggest the ability of the parasites to modulate the expression of the vector`s proteases. However, this modulation involved both down-regulation (e.g. LuloChym1A, a chymotrypsin and LuloCpepA1, a carboxypeptidase in *L. longipalpis*) and up-regulation (e.g. LlTryp2, a trypsin in *L. longipalpis*). Moreover, the presence of *Leishmania* also appears to modulate the abundance of proteases transcripts expressed after the blood meal has been digested (e.g. LuloTryp3, a trypsin in *L. longipalpis*). More detailed studies at both transcriptomic and biochemical level would be needed to show how this modulation affects the resulting proteolytic activities in the midgut. Also, the mechanism of this modulation remains unknown.

Another way the parasites might affect the gut proteolytic activity has recently been suggested. Serine protease inhibitors (ISP) were found in *L. major*, with inhibitory effect over vertebrate macrophage serine proteases, such as neutrophil elastase, trypsin and chymotrypsin (Eschenlauer et al. 2009). The possibility of ISP having an effect on insect midgut proteases is currently under investigation in our laboratory.

Perhaps most importantly, *Leishmania* parasites also possess mechanisms that increase their resistance to proteolytic attack without inhibiting the overall proteolytic activity in the midgut. Pimenta et al. (1997) exposed *L. major* to lysates of blood-fed *P. papatasi* midguts *in vitro*. Fresh tissue amastigotes and fully differentiated promastigotes were relatively resistant to killing during this exposure, whereas parasites within early stage amastigote-to-promastigote transition (2–8 h) became highly susceptible to killing, with greater than 95% reduction in numbers of viable parasites. The authors interpret their observation as resistence of promastigotes as well as fully transformed promastigotes to the midgut proteolytic activity. It is in accordance with the findings of Dillon and Lane (1993b), who indicated that fully differentiated promastigotes thrive in the presence of trypsin in culture. The identification of molecules that might play a role in defending the parasite against proteolytic damage has focused on a family of glycoconjugates, the phosphoglycans (PG), that incorpororate the common structure of repeating [Gal-Man-PO_4] units. These molecules are either attached to the cell surface through glycosylphosphatidylinositol (GPI) lipid anchors, including the lipophosphoglycan (LPG) and the proteophosphoglycan (PPG),

or they are secreted as protein-containing phosphoglycans, including the secreted proteophosphoglycan (sPPG) and a secreted acid phosphatase (reviewed by Sacks and Kamhawi 2001). The results of Sacks et al. (2000) imply phosphoglycans other than LPG in conferring resistence to the midgut digestive enzymes.

Taken together, sand fly midgut proteolytic enzymes are one of the critical factors affecting *Leishmania* development in the vector and represent attractive targets for vector-based transmission blocking strategy. Examples have been set in the research of malaria, where promising vaccine candidates have been identified being able to block *Plasmodium* ookinete development in the mosquito midgut (Ramirez et al. 2009). *Anopheles gambiae* midgut carboxypeptidase B has been shown to be up-regulated by *Plasmodium* infection and antibodies against one of these enzymes blocked parasite development in the mosquito midgut (Lavazec et al. 2007). Similarly, antibodies targeting an *A. gambiae* membrane aminopeptidase disrupted the development of *Plamsodium falciparum* and *P. berghei* ookinetes (Dinglasan et al. 2007).

1.2.2 Peritrophic matrix proteins and chitinases

The peritrophic matrix (PM) is an extracellular chitin-containing envelope, which in most insects separates the gut lumen from the midgut epithelium. It is composed of chitin, proteins, and glycoproteins (Lehane 1997). In nematoceran Diptera, including sand flies, females produce a type 1 PM, which is secreted by the midgut epithelium in direct response to the distension of the midgut caused by blood feeding (Jacobs-Lorena and Oo 1996). The structure of the sand fly peritrophic matrix is complex and rearranges during the course of blood digestion. Within several hours post blood meal (PBM) a thin PM composed mainly of chitin fibrils covers the whole surface of the blood bolus. At later stages (12h-2 days PBM depending on the sand fly species) the PM gets thicker and matures. Proteins and glycoproteins are incorporated in its structure and heme incrustations also appear. Sequentially (2-3 days PBM), the PM structure appears wrinkled to folded and then starts to break down (Secundino et al. 2005, Sadlova and Volf 2009).

At the molecular level, only a handful of proteins participating on the PM formation and breakdown in sand flies have been identified. The main class includes peritrophins, proteins contatining chitin binding domains (CBDs). Two types of putative peritrophin molecules have been identified in the midgut transcriptomes of *P. papatasi* and *L. longipalpis*: multiple-CBD peritrophins (PpPer1, 4 domains; PpPer3, 3 domains;

LuloPer1, 4 domains) and single-CBD proteins (PpPer2, LuloPer2, and LuloPer3) (Ramalho-Ortigao et al. 2007b, Jochim et al. 2008). Some of them contain predicted -N and/or -O type glycosylation sites including mucin-type domains. The putative peritrophins with multiple CBDs are likely to have a role in cross-linking the chitin fibrils of the peritrophic matrix. Single-CBD peritrophins may have roles in capping the ends of chitin fibrils or sequestering free chitinous molecules within the midgut lumen. In addition to chitin binding, mosquito proteins with CBDs have also been described to bind heme and have a role in its sequestration during blood digestion (Devenport et al. 2006).

Glycosylation of the PM proteins can be of great importance for the PM structure and function. Heavily glycosylated proteins can influence the selectiveness of the PM pores and account for water retention within the PM. Glycosylation can also influence susceptibility to degradation by temporally secreted digestive proteases thus affecting the changes in the PM thickness and structure. Interestingly, comparative transcriptomic studies have shown modulation of peritrophin transcripts abundance by the presence of *Leishmania* parasites. *Phlebotomus papatasi* infected with *L. majo*r downregulated the multi-domain peritrophin (PpPer1), whereas *L. longipalpis* infected with *L. infantum* upregulated the orthologous peritrophin (LuloPer1) (Ramalho-Ortigao et al. 2007b, Jochim et al. 2008). The significance of this finding remains unclear. Besides peritrophins, non-chitin binding proteins have recently been identified in the PM of *Anopheles gambiae* (Dinglasan et al. 2009). It is conceivable that some non-chitin binding proteins also participate on PM formation in sand flies.

The major role in PM breakdown has been attributed to chitinases. Chitinolytic activity in the sand fly midgut is induced after the intake of blood and peaks around 48h PBM. Ramalho-Ortigao et al. (2005) described a functional, blood-induced chitinolytic system, in the midgut of *P. papatasi* and named the identified enyzme PpChit1. It was produced as a recombinant protein and antibodies against this protein inhibit the midgut chitinolytic activity *in vitro*. The authors presume that PpChit1 is involved in the maturation and degradation of *P. papatasi* PM (and suggest the same role for its ortholog, LlChit1, in *L. longipalpis*) (Ramalho-Ortigao and Traub-Cseko 2003, Ramalho-Ortigao et al. 2005).

Several studies suggest a dual role for the sand fly PM regarding *Leishmania* development: it protects the parasites against proteolytic attack in the beginning of digestion yet becoming a barrier to parasite escape when mature. Pimenta et al. (1997)

studied the role of PM in *L. major* development in *P. papatasi*. They observed that blocking the PM formation by addition of chitinase in the bloodmeal leads to a sharp increase in the number of parasites killed within a few hours PBM. Early parasite mortality was reversed by the additon of soybean trypsin inhibitor. The authors conclude that the PM creates a barrier to the rapid diffusion of digestive enzymes, and limits the exposure of parasites to these enzymes during the time when they are especially vulnerable to proteolytic damage (Pimenta et al. 1997). On the other hand, at later phases the PM appears to act as a barrier to the parasite development. Long nectomonads must escape from the endoperitrophic space, through the PM, to prevent being passed together with remnants of the digested blood meal. Walters et al. (1992) report entrapment of *L. panamensis* in the endoperitrophic space of *P. papatasi*. The failure of the parasite to escape from the PM in an inappropriate vector resulted in their expulsion from the midgut. Pimenta et al. (1997) further showed that addition of allosamidin, a chitinase inhibitor, to infective blood meal led to thickening of the PM and entrapment of *L. major* within the peritrophic space thus preventing further development of the parasite in its natural vector *P. papatasi*. Recent data also indicate that an anterior PM plug located at the junction between the anterior and posterior midgut acts as a barrier to *Leishmania* migration towards the stomodeal valve (Sadlova and Volf 2009).

Schlein et al. (1991) first proposed that *Leishmania* escape from the PM is accomplished by a parasite chitinase. They described *L. major* escaping at the anterior end of the PM in *P. papatasi*. Further work supported their hypothesis by showing that *L. mexicana* chitinase-overexpressing strain had an accelerated escape from the PM in *L. longipalpis* (Rogers et al. 2008). However, in a recent study, Sadlova and Volf (2009) suggest that *L. major* chitinase does not have an important role in the disintegration of PM in *P. duboscqi*. The detailed histological and electron-microscope study did not reveal any signs of PM lysis caused by *Leishmania* and showed that the PM opens similarly in uninfected and infected females. *Leishmania major* parasites escaped from the posterior end of PM opened at the end of blood meal digestion (lysed presumably by the activity of vector chitinase). Importantly, Coutinho-Abreu et al. (2010) report that knock-down of *P. papatasi* chitinase, PpChit1, by the means of RNAi led to a significant reduction in the number of *L. major* present in the midgut 120h PBM. It can be concluded that the parasites taking advantage of the sand fly chitinolytic activity within the midgut is the main mechanism for their escape. Taken together, the PM plays

important roles in the parasite development and proteins involved in its formation, maturation and desintegration provide a promising target for transmission blocking vaccines.

1.2.3 Proteins and peptides involved in innate immunity

Innate immune response plays a role in the control of bacterial and parasitic infections in the midgut of bloodsucking insects (Hao et al. 2001, Feldhaar and Gross 2008). In sand flies, defensins, cationic antibacterial peptides, have been described in the fat body and the midgut. In *P. duboscqi*, defensin was induced by both bacteria and *Leishmania* infection and the recombinant peptide showed a significant antiparasitric activity against *L. major in vitro* (Boulanger et al. 2004). Transcripts coding for several other putative components of the innate immune response have been detected in the sand fly midgut, such as pattern recognition proteins, a glycin-rich protein and serpins (Ramalho-Ortigao et al. 2007b, Jochim et al. 2008, Pitaluga et al. 2009). The role of these proteins and that of the sand fly immunity in general in *Leishmania* infection await further investigation.

1.3 Attachment of the parasites to the midgut epithelium

Following the escape from the endoperitrophic space, the parasites attach to the midgut, inserting their flagella between the epithelial microvilli. There is no obvious ultrastructural modification of the flagellum associated with midgut binding. It remains unclear whether the involvement of the flagellum is essential *per se*, or merely a reflection that being at the anterior end the flagellum will contact the epithelium first and fits between the microvilli (reviewed by Bates 2008). By anchoring themselves to the midgut the parasites help to prevent their expulsion from the gut during defecation, and it has been postulated that this binding is the main determinant of parasite-vector specificity (Pimenta et al. 1994, Sacks 2000).

Based upon experimental tests of their ability to support development of wide or limited range of *Leishmania* species, sand flies have been classified as permissive or specific (also called restrictive by some authors) vectors (Volf and Myskova 2007).

Most sand fly species tested to date support development of multiple *Leishmania* species and are thus called 'permissive vectors'. In contrast, there appears to be a close evolutionary fit between *P. papatasi* and *P. sergenti* with *L. major* and *L. tropica* respectively, as other *Leishmania* species survive poorly in these sand fly hosts reviewed by (Kamhawi 2006). The mechanism of parasite attachment has been most intensively studied in the specific vector *P. papatasi* infected with *L. major*.

1.3.1 Phlebotomus papatasi

The attachment of *Leishmania major* in its specific vector *P. papatasi* is perhaps the best characterized parasite-sand fly interaction so far. The role of parasite surface lipophosphoglycan (LPG), has been demonstrated by a series of studies. LPG is an abundant glycolipid that covers the entire surface, including the flagellum, of all *Leishmania* promastigote stages. The basic LPG structure is highly conserved in all *Leishmania* species. It consists of a glycosyl-phosphatidyl-inositol lipid anchor attached through a hexasacharide core to a polymer of 10–30 PG repeating units terminated by a small neutral oligosaccharide cap (Turco and Descoteaux 1992). The PG repeating units are often modified by strain-, species-, and stage-specific side-chain sugar residues. Purified *L. major* LPG was shown to bind to dissected *P. papatasi* midguts (Lang et al. 1991) and inhibit the binding of *L. major* promastigotes to the midgut *in vitro* (Pimenta et al. 1992). More recent studies using LPG-defecient parasites confirmed the crucial role of LPG in the attachment of *L. major* in the midgut. These mutants lack the LPG1 gene which encodes a galactofuranosyltransferase required for synthesis of the LPG glycan core, rendering such cells specifically deficient in LPG. The ability to persist in the midgut of *P. papatasi* following bloodmeal excretion was completely lost in these parasites and this defect was correlated with their inability to bind to midgut epithelial cells *in vitro* (Sacks et al. 2000).

In order to produce a transmissible infection in the sand fly, the parasites need to be able to detach from the midgut epithelium and produce free-swimming metacyclic forms. In *L. major - P. papatasi* combination the attachment is achieved by stage-specific modifications in the LPG structure. Parasite binding is mediated by modified phosphoglycan repeats bearing side chain galactosyl residues (Pimenta et al. 1992). During metacyclogenesis, the original LPG is shed and replaced by metacyclic form LPG, which has increased numbers of PG repeats and side-chaingalactose residues

masked by the addition of terminal arabinose (Mcconville et al. 1992). Thus modified metacyclic form LPG does no longer bind to the *P. papatasi* midgut (Pimenta et al. 1992).

Based on the finding of the role of sugar residues in the attachment a hypothesis was postulated that lectins or lectin-like molecules serve as receptors for parasite binding in the midgut. Lectin-like activities have indeed been described in the sand fly midgut (Wallbanks et al. 1986, Palanova and Volf 1997, Volf et al. 1995, Palanova and Volf 1997). Sequencing of a *P. papatasi* midgut cDNA library led to the discovery of a galectin molecule (PpGalec) that was proved to serve as a receptor for *L. major* LPG (Kamhawi et al. 2004). PpGalec is a 35kDa galectin containing two non-identical carbohydrate recognition domains. It is expressed throughout the development of larval and pupal stages, but is strongly up-regulated in adult females. Its expression appears to be restricted to the midgut. Despite lacking a signal peptide, it is expressed on the luminal surface of *P. papatasi* midgut epithelial cells. This is consistent with the cell surface localization of some other galectins, that are known to be trafficked through non-classical secretory pathways to the cell surface, where they bind to appropriately glycosylated surface molecules (Rabinovich and Gruppi 2005).

The role of PpGalec in *L. major* binding was proved by several experiments. PpGalec produced as a recombinant protein bound specifically to *L. major* promastigotes bearing side-chains galactose residues on their LPG *in vitro*. Antibodies raised against this protein blocked *L. major* binding to midguts *in vitro* and and severely impaired the parasites` development *in-vivo* when fed to *P. papatasi* in the infectious blood meal (Kamhawi et al. 2004). Interestingly, the binding of recombinant PpGalec to promastigotes was not only species-specific (recognizing neither *L. tropica* nor *L. donovani*), but also strain-restricted. Significant binding was only observed with the Friedlin V1 strain of *L. major* (Israeli isolate), sympatric to the *P. papatasi* used in the study (a colony originating from the Jordan Valley). A West African Seidman strain (SD) of *L. major*, with LPG virtually devoid of galactose side-chains as well the LV39 strain (Central Asia isolate), with long poly-galactose side chains failed to bind recombinant PpGalec (of the Jordanian flies). These observations are in accordance with the earlier findings that the LV39 grow poorly in the Cyprus as well Jordanian colonies of *P. papatasi* (Cihakova and Volf 1997, Sacks et al. 2000) and the SD strain does not survive in *P. papatasi* at all (Joshi et al. 1998). Using *L. major* lines mutated in galactosytransferases, Dobson et al. (2010) have recently characterized an LPG side-

chain galactosylation pattern optimal for survival in *P. papatasi* originating from the Jordan Valley. The key element is the presence of mostly mono-galactosylated PG repeats. However, the study also reveals that the optimal galactosylation pattern, while being a prerequisite, is not on its own sufficient for the binding to occur. Coating *L. donovani* with the optimally galactosylated LPG did not confer its survival in the midgut of *P. papatasi*, parasites being lost most likely due to the failure to bind to the epithelia. The authors suggest the existence of an additional, as yet uncharacterized *L. major*-specific ligand that is required for successful binding and survival in the midgut. Whether this additional ligand binding could also explain the fact that attachment usually occurs via the flagellum, whereas LPG is found over the whole surface of promastigotes, remains to be elucidated. Thus, in spite of being the best characterized *Leishmania*-sand fly interaction, the binding of *L. major* in its natural vector *P. papata*si is not yet fully understood.

1.3.2 Other sand fly species

While the role of LPG in the attachment of *L. major* in *P. papatasi* has been unambiguously proved, the necessity of LPG on the parasite surface and the nature of receptors for parasite binding in the midgut are still in question in other sand fly species. The structure of LPG side-chains is highly species- and in some cases strain-specific. Similarly to *L. major*, the structure of LPG is different in metacyclic parasites than in other forms in other *Leishmania* species. For example, in an Indian strain of *L. donovani* the PG repeats are modified with glucose and this modification is down-regulated during metacyclogenesis, along with increasing the length of the PG backbone (Mahoney et al. 1999). In contrast, in a Sudanese strain of this species there are no side chains modifications at all (Thomas et al. 1992).

After the identification of *L. major* LPG as the parasite ligand for binding in *P. papatasi* midgut, a number of studies have been carried out implicating a similar role for LPG in other *Leishmania* species. Pimenta et al. (1994) observed binding of purified LPG from several *Leishmania* species (*L. donovani*, *L. major*, *L. amazonensis*) to the midguts of *P. argentipes in vitro*, corresponding to the ability of the parasites to survive in *P. argentipes* in laboratory infections. In contrast, *P. papatasi* midguts were only stained with LPG purified from *L. major*. A similarly high specificity was found for *L. tropica* in its vector *P. sergenti*; midguts were intensely stained following incubation with purified PG from *L. tropica* compared with PGs from *L. major* or *L.*

donovani (Kamhawi et al. 2000). Soares et al. (2002) blocked the binding of *L. infantum* to dissected midguts of its natural vector *L. longipalpis* by purified PG of this species. Moreover, Pimenta et al. (1994) also report that *L. donovani* R2D2 mutant deficient in LPG biosynthesis failed to survive beyond defecation in *P. argentipes*, suggesting the requirement for LPG in the binding. However, the R2D2 was obtained following heavy mutagenesis and selection for LPG deficiency (King and Turco 1988), leaving the possibility that non-specific deleterious effects account for the observed phenotype. Restoration of the LPG1 gene expression to R2D2 only weakly restored survival in *P. argentipes* (Sacks et al. 2000).

Despite the large body of evidence implying LPG in the parasite attachment, neither receptors in the sand fly midgut nor the mechamism of parasite release in the later phase of infection have been sufficiently characterized in sand flies other than *P. papatasi*. Expression of the tandem repeat galectin (PpGalec) seems to be restricted to *P. papatasi* and *P. duboscqi*, as shown by a genomic dot blot as well as immunoblot (using antisera raised against this protein) with a variety of sand fly species (Kamhawi et al. 2004). While its role in *L. major* binding in *P. papatasi* has been proved (see above), its function in *P. duboscqi* remains speculative. First, experimental infections of *P. duboscqi* with *L. major* LPG deficient mutants have only been performed with a small number of sand flies bringing rather ambiguous results as regards the necessity of LPG for the binding. The infections with *lpg1*$^{-}$ mutants persisted beyond defecation but their intensity decreased over time (Boulanger et al. 2004). Second, it is not clear how would the galectin bind the LPG of the SD strain of *L. major,* that develops mature infections in *P. duboscqi* (Joshi et al. 2002). As mentioned above, the LPG of this strain is devoid of any galactose side chains (Sacks et al. 1990).

In the analysis of the *L. longipalpis* midgut-specific transcriptome, only one low-abundance transcript was identified, which is homologous to a single-domain galectin (Jochim et al. 2008). It is unlikely that this galectin acts as a receptor for *Leishmania* in *L. longipalpis*. No molecular data are available for other sand fly species and the nature of putative LPG receptors remains unclear. Pimenta et al. (1994) suggested that midguts of *P. argentipes* possess a receptor for a conserved part of LPG, accounting for the broad permissivity to various *Leishmani*a species. In their later study, the authors suggest that the binding of *L. donovani* in *P. argentipes* occurs via receptors for sacharides present in the neutral LPG cap that is masked by conformational changes in the elongated PG chains in metacyclic *L. donovani* (Sacks et al. 1995). It should be

noted that despite sharing some common features (all are composed of neutral hexoses), the LPG caps show remarkable interspecies differences. They vary both quantitatively and qualitatively in the content of mannose, galactose or glucose, raising uncertainty about the nature of a putative common receptor.

Importantly, Myskova et al. (2007) have recently observed LPG-independent development of *Leishmania* in two permissive vectors, *P. arabicus* and *L. longipalpis*. *Leishmania major lpg1*$^{-}$ line devoid of LPG survived well and developed mature infections fully comparable to wild type parasites in these sand flies. Similarly, (Rogers et al. 2004) report that *L. mexicana lpg1*$^{-}$ mutants survive and complete their development in *L. longipalpis*. Myskova et al. (2007) hypothesised that LPG is required in specific vectors, while in permissive vectors *Leishmania* bind via an LPG independent mechanism. They observed a correlation between the occurence of N-acetyl-D-galactosamine- displaying glycoconjugates in the midgut of sand flies and their permissivity. They suggest these glycoconjugates as ligands for *Leishmania* attachment in permissive vectors and show the binding of such molecules in the midgut lysate of *P. halepensis* to *L. major* promastigotes *in vitro*. This new binding modality implies involvement of a parasite lectin-like receptor. The authors propose heparin binding proteins that had been previously described on the surface of various *Leishmania* species (Mukhopadhyay et al. 1989, Kock et al. 1997, Svobodova et al. 1997) as potential candidates. It should be noted that Myskova et al. (2007) regard the distinction of sand flies into specific and permissive vectors as a working concept that likely oversimplifies the real situation. They stress the need of further studies using different parasite-vector combinations.

In conclusion, it is clear that much still remains to be learned about the mechanisms of attachment on both the parasite and the vector side. More studies providing insights in the repertoire of sand fly midgut-expressed molecules, where the available data are scarce, are particularly desirable.

2 Objectives

The PhD project was aimed at characterizing *Leishmania* attachment in sand fly midgut and at the role of sand fly midgut proteins in the interactions with the parasites. I sought to identify *Leishmania* life-cycle stages binding to the midgut epithelium and to assess the species-specifity of the attachment. I also investigated the role of the parasite`s surface glycoconjugates in the attachment and development in the midgut of several sand fly species. Last but not least, I was interested in the proteins expressed in the midgut of *Phlebotomus perniciosus*, a permissive vector of leishmaniasis, and their possible roles in the interactions with the parasite.

The main objectives of this thesis were:

- to determine the stage- and species-specificity of *Leishmania* binding to the sand fly midgut using a novel comparative *in vitro* binding assay

- to test the role of the surface lipophospoglycan and other phosphoglycans in *Leishmania major* development in its natural vector *Phlebotomus duboscqi* and two permissive sand fly species, *Phlebotomus perniciosus* and *Phlebotomus argentipes*

- to generate, annotate and analyze midgut-specific cDNA libraries of *P. perniciosus*

3 Publications

Wilson R, Bates MD, **Dostálová A**, Ječná L, Dillon RJ, Volf P, Bates PA. Stage-specific adhesion of *Leishmania* promastigotes to sand fly midguts assessed using an improved comparative binding assay. PLoS Neglected Tropical Diseases, 2010, 4(9):e816. PMID: 20838647

Svárovská A, Ant TH, Šeblová V, Ječná L, Beverley SM, Volf P. *Leishmania major* glycosylation mutants require phosphoglycans (lpg2-) but not lipophosphoglycan (lpg1-) for survival in permissive sand fly vectors. PLoS Neglected Tropical Diseases, 2010, 4(1):e580. PMID: 20084096

Dostálová A, Favreau AJ, Barbian KD, Votýpka J, Volf P, Valenzuela JG, Jochim RC. The midgut transcriptome of *Phlebotomus* (*Larroussius*) *perniciosus*, a vector of *Leishmania infantum*: comparison of sugar fed and blood fed sand flies. Manuscript accepted for publication in BMC Genomics.

4 Summary and conclusions

This dissertation thesis sums up published results of three projects I participated in during my PhD studies. The projects were focused on *Leishmania*-sand fly interactions during the parasite development in the midgut. Main emphasis was put on permissive vectors, the knowledge of which had in many ways lagged behind the best characterized vector *Phlebotomus papatasi*, highly specific for *Leishmania major*.

First, we investigated the stage- and species-specificity of *Leishmania* attachment in the midgut. To this end, we developed a comparative *in vitro* binding assay, which allows examination of two differentially stained parasite populations bound to the same midgut. This relative binding methodology overcomes difficulties posed by the large variation in numbers of parasites bound to individual midguts, which was noted in earlier studies (Pimenta et al. 1994, Kamhawi et al. 2000). As opposed to the previous assays using PBS, we incubated the guts in Grace`s insect medium, providing greater stability of the midgut tissue. This also allowed us to increase the incubation time with the parasite. Moreover, we used midguts of flies that had recently passed the bloodmeal rather than unfed flies used in earlier studies. Overall, we achieved more physiological conditions better mimicking the situation *in vivo.*

To determine the life-cycle stages involved in the binding, we prepared promastigote population of *L. mexicana* and *L. infantum* enriched for the four major developmental forms: procyclic promastigotes, long nectomonads, short nectomonads and metacyclic promastigotes. Our results show that both long and short nectomonads of both *Leishmania* species bind in high numbers to the midguts of *L. longipalpis*. Direct comparison of the two stages showed that short nectomonads bind slightly less efficiently in both *Leishmania* species, though the difference was statistically significant only in *L. infantum*. In contrast, there was very little binding of either procyclic or metacyclic promastigote-enriched populations in either of the two *Leishmania* species. From these data we can infer that long nectomonad promastigotes constitute the first population of parasites bound to sand fly midgut during parasite development *in vivo*. Eventually, they are replaced by the short nectomonads, which are the next developmental form and prevail in the midgut from day 4 post infection onwards (Gossage et al. 2003). However, the predominant form bound to the sand fly gut at a

particular point of development will depend in great part on the detachment rate, which is currently unknown.

The observed inability of metacyclic promastigotes to bind is in accordance with the results of earlier studies, which mostly attributed this loss in the binding capacity to changes in the structure of LPG during metacyclogenesis (Sacks 2001). However, these modifications are insufficient by themselves to explain lack of metacyclic binding in *L. longipalpis*, where LPG-independent attachment mechanisms have been described (Myskova et al. 2007).

Procyclic promastigotes, on the other hand, may be unable to bind to the midgut due to a lack of (sufficient) LPG on their surface, since *in situ* immunohistological studies in *L. major* failed to detect any surface LPG until day 3 post-blood meal, a time point which corresponding with the appearance of the long nectomonad promastigotes (Saraiva et al. 1995). Alternatively, complying with the LPG-independent attachment theory, there is the possibility of procyclic promastigotes being unable to bind to the midgut for mechanical reasons. First, having a very short flagellum, procyclic parasites are only weakly motile. Second, the flagellum length may be insufficient for it to mediate binding to sand fly receptors, were these restricted to the base of microvilli.

Following the studies of stage-specificity, we then went on to determine the ability of various *Leishmania* species to bind to the midguts of both specific and permissive vectors *in vitro*. In our experiments with *L. longipalpis*, all the tested *Leishmani*a species were able to bind to the midgut in significant numbers, which is in accordance with previous observations (Pinto-da-Silva et al. 2005). However, surprising results were obtained through direct comparison of the binding of *L. infantum*, the parasite species that is transmitted by *L. longipalpis* in nature, to other species. In line with expectations, *L. infantum* was able to bind more efficiently than either *L. mexicana* or *L. braziliensis*. On the other hand, *L. major* and *L. tropica* both strongly out-competed *L. infantum*. It is not clear whether the parasites are binding to the same receptors, with *L. major* and *L. tropica* simply having a higher affinity, or each species is binding to distinct receptors. In the latter scernario, the binding of one species is much stronger and steric hindrance is preventing the other from accessing their receptor. Either way, these results show that the natural parasite is not necessarily the species that can always bind *in vitro* most effectively to the midgut of its vector.

In the specific vector, *P. papatasi*, we were surprised to find that both *L. braziliensis* and *L. tropica* were able to bind to the midgut in significant numbers when

competing with the natural parasite *L. major*. Neither of these species is able to complete their development in *P. papatasi in vivo*. An observation similar to ours was made in studies on the flagellar attachment of *L. panamensis*, a peripylarian species closely related to *L. braziliensis* (Warburg et al. 1989). In contrast, the observed binding of *L. tropica* is contradictory to the results of Pimenta et al. (1994), who observed poor binding of *L. tropica* to the midguts of *P. papatasi*.

We also tested binding of *L. major* and *L. tropica* to midguts of another sand fly species, *P. sergenti*. This is the specific vector of *L. tropica* and, therefore, this parasite might be expected to out-compete all other *Leishmania* species. However, there was equal *in vitro* binding of *L. major* and *L. tropica* in *P. sergenti.* These results contradict those reported by Kamhawi et al. (2000), who reported only very small numbers of *L. major* bound to *P. sergenti* midguts as compared to midguts incubated with *L. tropica*. We suggest, that these apparently conflicting results may have arisen from different parasite strains and sand fly colonies originating from different areas being used or from differences in the expermintal design of the binding assays.

In order to elucidate molecular mechanisms underlying the binding and development of *Leishmania* parasites in the midgut of permissive vectors, we tested the role of *Leishmania* surface phosphoglycans in three sand fly species: *P. perniciosus*, *P. argentipes* and *P. duboscqi.* Glycolipids an glycoproteins containing a conserved [Gal-Man-PO4] phosphoglycan (PG) unit belong to the most abundant surface-localized and secreted molecules synthesised by *Leishmania* promastigotes. As discussed in the introduction to this thesis, these molecules have been implicated in a number of functions in the development in the sand fly midgut including protection againts proteolytic attack, binding to the midgut and blocking the stomodeal valve (Kamhawi 2006).

First we looked into the role of the surface lipophosphoglycan (LPG) in *Leishmania* development. The necessity of LPG for the midgut attachment of *L. major* in its specific vector *P. papatasi* is well documented (Sacks et al. 2000). On the other hand, LPG-independent development of *Leishmania* has been recently described in two permissive vectors, *L. longipalpis* and *P. arabicus* (Myskova et al. 2007). Knowing that *P. perniciosus* and *P. argentipes* are broadly permissive to development of various *Leishmania* species in laboratory infections, we speculated that they might be able to accommodate the parasite without the requirement of LPG on its surface. We tested this

hypothesis using *L. major* null mutants in the LPG1 gene, which encodes a galactofuranosyl transferase involved in biosynthesis of the LPG glycan core. We found that this line specifically lacking LPG remains able to develop in the permissive vectors *P. perniciosus* and *P. argentipes* at levels resembling those of wild type parasites, with full midgut development and colonization of the stomodeal valve. Our results contradict those reported by Pimenta et al. (1994). In their study, *L. donovani* LPG-deficient mutant line 'R2D2' failed to survive in *P. argentipes*. We suggest that the effect they observed can be attributed to non-specific deleterious effects of heavy mutagenesis used in generation of the 'R2D2' line, as restoration of LPG1 expression to this line only weakly restored survival in *P. argentipes* (Sacks et al. 2000).

Another sand fly species that we tested is *P. duboscqi*, a vector of *L. major* in sub-Saharan Africa (Beach et al. 1984). The degree of permissivity of this species was not clear. Some populations have been shown experimentally to support development of *L. tropica* to a certain extent (Killick-Kendrick et al. 1994) but development of other *Leishmania* species had not been tested. To address this question, we infected *P. duboscqi* with *L. infantum*. We observed complete elimination of the parasites by day 8 post infection, demonstrating that *P. duboscqi* is refractory to *L. infantum* and is therefore classified as a restrictive or specific vector. We went on to test the LPG deficient mutants development in this sand fly species. We found that in contrast to the permissive vectors, the development of these mutants was severely impaired in *P. duboscqi*. Although the early infections were similar to those of the wild type parasites, there was a substantial decrease in the infections rate after defecation of the bloodmeal. The LPG deficient parasites produced mature late stage infections in very few females only.

Our results extend those reported in the study by Boulanger et al. (2004) performed with a small number of sand flies. Similar results with *L. major* LPG deficient mutants in *P. duboscqi* were obtained also by Secundino et al. (2010). Taken together, our data argue that, *Leishmania* LPG is required for late phase survival in the specific vector *P. duboscqi* but not in the permissive sand flies *P. argentipes* and *P. perniciosus*. These results correlate with the presence of midgut glycoproteins bearing terminal N-acetyl-galactosamine, that were previously suggested as potential ligands for *Leishmania* binding via an LPG-independent mechanism (Myskova et al. 2007). While present in both permissive species tested (Myskova et al. 2007), we report that such glycoconjugates are absent in *P. duboscqi* midgut.

Using *L. major lpg2*$^-$ mutants, we also studied the role of phospoglycans other than LPG in the development in the three sand fly species. These parasites lack the LPG2 gene encoding a Golgi GDP-sugar transporter, resulting in failure to synthesize not only LPG but also all other phospoglycans (PGs) (Spath et al. 2003). We observed that parasites devoid of PGs were impaired in early development in sand fly and unable to survive in all sand fly species tested. In *P. argentipes* and *P. duboscqi* parasites of this line were severely impaired as early as day 2 post infection. For *P. duboscqi*, these data stand in line with those of Boulanger et al. (2004). A similar finding was described for PG-deficient mutants of *L. donovani* in *P. argentipes* (Sacks et al. 2000). Moreover, simultaneously to our study, Secundino et al. (2010) made similar observations in *P. duboscqi* originating from Mali and in *L. longipalpis*.

It has been hypothesized that parasite death in the predefecated sand fly midgut is attributable to digestive enzymes and that phosphoglycans other than LPG confer resistance to the proteolytic attack (Pimenta et al. 1997, Sacks et al. 2000). In our study, we assessed the importance of LPG2-dependent molecules by exposing parasites to the activity of a proteolytic enzyme *in vitro*. As trypsin-like proteases were described as the most abundant digestive enzymes in both *P. papatasi* and *L. longipalpis* midgut after bloodfeeding (Ramalho-Ortigao et al. 2007b, Jochim et al. 2008), bovine panceratic trypsin was chosen for these experiments. We did not observe any differences in the growth of wild type and mutated parasites when exposed to high concentrations of trypsin. Thus, the *lpg2*$^-$ promastigotes lacking surface PGs did not prove to be more vulnerable to trypsin activity than the wild type. These results could argue against a role of PGs in conferring resistance of promastigotes to the trypsin-like digestive enzymes in the sand fly gut. However, in light of the studies of Secundino et al. (2010), published simultaneously with our work, we think it more likely that bovine trypsin is not a good model for the activity and/or properties of all the proteolytic contents of the sand fly midgut. Secundino et al. showed that addition of soybean trypsin inhibitor in the bloodmeal does in part rescue the survival of *lpg2*$^-$ parasites in the midgut. They identified proteophosphoglycan (PPG), likely its secreted form, as the key molecule conferring resistence to the midgut digestive enyzmes. Addition of purified PPG prevented in great part the killing of promastigotes exposed to midgut lysates prepared from blood-fed flies *in vitro*. Protection was not associated with inhibition of enzyme activities, but with cell surface acquisition of the PPG (Secundino et al. 2010). However, as inhibition of trypsin actvity (as well as general proteolytic activity) only

provided partial rescue to the *lpg2*$^{-}$ parasites, the authors do not exclude the possibility that also factors other than proteolytic attack contribute to the inability of *lpg2*$^{-}$ parasites to survive within the bloodmeal.

In the third study, we focused on molecules expressed in the midgut of the permissive vector, *P. perniciosus*, which are most likely to have important interactions the transmitted parasites. We constructed, sequenced and analyzed two midgut specific cDNA libraries from sugar fed and blood fed female *P. perniciosus* and compared the transcript expression profiles. We have provided a catalogue of molecules with putative roles in the blood meal digestion, peritrophic matrix formation, immunity and response to oxidative stress. By detailed comparison of our findings with the published midgut transcriptomes of other sand fly species, *L. longipalpis* and *P. papatasi* (Ramalho-Ortigao et al. 2007b, Jochim et al. 2008, Pitaluga et al. 2009), we identified several features shared by the two permissive vectors, *P. perniciosus* and *L. longipalpis*, potentially important for their vectorial competence. We obtained a total of 4111 high quality sequences and grouped them by homology into 370 contigs and 1085 singletons in a combined analysis of the two libraries.

Proteolytic enzymes were among the most abundant transcripts we detected. These included putative trypsins, chymotrypsins, metallo-carboxypeptidases, astacins and an alanyl-aminopeptidase. Interestingly, two of these putative enyzmes (PperChym5 and PperAstacin2) are novel molecules distantly related to putative chymotrypsins and astacins, respectively, described in other sand flies. Similarly to *P. papatasi* and *L. longipalpis*, some of the putative proteases seem to be up-regulated after blood feeding (PperChym1, PperChym2 and PperTryp3), while the abundance of other decreases after the intake of blood (PperTryp1, PperTryp2 and PperChym3). The latter three likely represent post-transcriptionally regulated "early" digestive proteases that are stored in the midgut prior to blood feeding, similarly as described in some mosquitoes (Graf and Briegel 1989, Barillas-Mury et al. 1995).

We performed a temporal expression profile analysis of the three putative trypsin transcripts by the means of quantitative PCR. PperTryp1, the most abundant trypsin identified, was downregulated as soon as 6h after blood feeding, further suppressed 24h post-blood meal and returned to the pre-blood meal levels in the sand flies that had passed the remnants of blood meal. PperTryp2 was detected in lower amounts (about 1/70th of PperTryp1) and represents another trypsin down regulated by

blood feeding. On the contrary, PperTryp3 was confirmed as the main blood feeding-induced trypsin molecule, elevated after 6h and reaching the highest quantity 24h post-blood meal. Interestingly, its abundance at this time point was decreased in the presence of *L. infantum*, suggesting the ability of the parasite to suppress or delay its expression. Our finding is in accordance with a decrease or delay in midgut trypsin activity in sand flies infected with *Leishmania* reported in several studies (Schlein and Romano 1986, Dillon and Lane 1993b, Sant'Anna et al. 2009). Notably, Telleria et al. (2010) have recently made such an observation in *L. longipalpis* infected with *L. infantum*. However, it should be noted that the extent to which the obsereved modest decrease in PperTryp3 transcript abundance affects the overall midgut trypsin activity is not known and neither is the mechanism by which *Leishmania* could influence the expression of sand fly molecules.

Another group of putative transcripts we identified code for proteins putatively involved in peritrophic matrix (PM) formation, maturation and decomposition. We described three putative peritrophins containing chitin binding domains. Most of the sequences originated from the library of midguts after blood feeding, which is in accordance with putative peritrophins sequence distribution in *L. longipalpis* (Jochim et al. 2008). Interestingly, different peritrophin expression pattern was described in specific vector *P. papatasi* where a peritrophin transcript was highly abundant before blood feeding and thereafter down-regulated (Ramalho-Ortigao et al. 2007b). Besides peritrophins, we also identified two transcripts homologous to proteins that have been recently identified in the peritrophic matrix of *Anopehles gambiae* (Dinglasan et al. 2009). We found homologs of these proteins among the *L. longipalpis* and *P. papatasi* midgut transcripts and suggest that these proteins might also participate in the PM formation in sand flies. Furthermore, we describe a putative *P. perniciosus* midgut chitinase detected in the library from midguts after blood feeding. The identified proteins putatively affecting structure, permeability and rate of formation/decompsosition of the PM are of special interest given the important roles the PM plays in the early phase of *Leishmania* development in the midgut.

The overall most abundant sequences belong to a group of transcripts coding for putative proteins containing a conserved insect-allergen domain. We called these microvillar proteins (MVPs) for their similarity to mosquito protein of unknown function (named G12) localized on microvillar membranes of midgut epithelial cells (Shao et al. 2005). We identified five putative MVPs in *P. perniciosus*. Apart from the

insect-allergen domain, these five putative proteins share little sequence homology suggesting they may have different functions altogether. Five respective MVPs homologs were previously described in the midgut of *L. longipalpis* and four in *P. papatasi*. Interestingly, no sequence with high similarity to PperMVP5 was found in the midgut of the specific vector *P. papatasi*. The MVPs seem to be induced only after blood feeding with the exception of PperMVP3 that is in contrast down-regulated by bloodmeal intake. Functions of these highly abundant, strictly regulated proteins remains yet to be elucidated.

Importantly, we also identified a number of molecules putatively involved in immunity and protection against oxidative stress in the midgut. Among these, there were several novel molecules, that had not been previously described in the midgut of any sand fly species. Novel putative antioxidant enzymes include an intracellular superoxide dismutase (PperSOD2) and putative microsomal and Theta class glutathione S-transferases (PperGST2 and PperGST3). Redox-related molecules do not only protect the midgut from oxidative damage, but were also shown to regulate midgut epithelial immunity and impact the outcome of bacterial and parasitic infections in mosquitoes (Kumar et al. 2003, Jaramillo-Gutierrez et al. 2009, Kumar et al. 2010). We assume a similar role of antioxidative enzymes in the sand fly midgut. We also describe novel molecules potentially involved in pathogen recognition such as a gram-negative bacteria binding protein (PperGNBP) and a putative peptidoglycan receptor (PperPGRPLC). Pathogen pattern recognition proteins play central roles in activating insect immune reactions and homologues of the identified *P. perniciosus* proteins have been shown to be involved in mosquito defence against bacteria and *Plasmodium* infections (Warr et al. 2008, Meister et al. 2009). It is likely that the identified *P. perniciosus* pattern recognition proteins play a role in protection against bacteria and possibly also *Leishmania* infection in the midgut.

Altogether, the studies summarized in this thesis have brougth interesting findings about several crucial points during *Leishmania* development in the sand fly midgut. First, we demonstrated that *Leishmania* gut binding is strictly stage-dependent, is a property of forms found in the middle phase of development (long nectomonad and short nectomonad forms), but is absent in early forms occurring in within the blood meal, procyclics, and in final stages, metacyclics. We confirmed the essential role of parasite surface LPG in the specific vector *P. duboscqi*, while we demonstrated the

existence of an LPG-independent mechanism of attachment in *P. perniciosus* and *P. argentipes*, as a feature common to permissive sand fly species. Furthermore we showed that although *Leishmania* binding to sand fly midgut may be necessary for parasite establishment, the specificity of *in vitro* binding alone is insufficient to explain vector specificity and other factors must play a role in limiting the development of parasites in certain vectors. Lastly, we provided an insight into the molecular background of events that occur in the midgut of *P. perniciosus*. By detailed comparison to two other sand flies species differing in their abitily to support the development of *Leishmania* we identify differences and similarities potentially important for the vectorial competence of sand flies and provide new targets for vector control.

5 References

Adler, S. 1938. Factors determining the behaviour of Leishmania sp. in sandflies. Harefuah 14.

Adler, S., and O. Theodor 1935. Investigations on Mediterranean Kala Azar. IX. Feeding Experiments with Phlebotomus perniciosus and other Species on Animals Infected with Leishmania infantum. Proc. R. Soc. B 116:516-542.

Barillas-Mury, C. V., F. G. Noriega, and M. A. Wells 1995. Early Trypsin Activity is Part of the Signal Transduction System that Activates Transcription of the Late Trypsin Gene in the Midgut of the Mosquito, Aedes aegypti. Insect Biochem. Mol. Biol. 25:241-246.

Bates, P. A. 2007. Transmission of Leishmania metacyclic promastigotes by phlebotomine sand flies. Int. J. Parasitol. 37:1097-1106.

Bates, P. A. 2008. Leishmania sand fly interaction: progress and challenges. Curr. Opin. Microbiol. 11:340-344.

Beach, R., G. Kiilu, L. Hendricks, C. Oster, and J. Leeuwenburg 1984. Cutaneous leishmaniasis in Kenya - transmission of Leishmania major to man by the bite of a naturally infected Phlebotomus duboscqi. Trans. Roy. Soc. Trop. Med. Hyg. 78:747-751.

Borovsky, D., and Y. Schlein 1987. Trypsin and chymotrypsin-like enzymes of the sandfly Phlebotomus papatasi infected with Leishmania and their possible role in vector competence. Med. Vet. Entomol. 1:235-242.

Boulanger, N., C. Lowenberger, P. Volf, R. Ursic, L. Sigutova, L. Sabatier, M. Svobodova, S. M. Beverley, G. Spath, R. Brun, B. Pesson, and P. Bulet 2004. Characterization of a defensin from the sand fly Phlebotomus duboscqi induced by challenge with bacteria or the protozoan parasite Leishmania major. Infect. Immun. 72:7140-7146.

Cihakova, J., and P. Volf 1997. Development of different Leishmania major strains in the vector sandflies Phlebotomus papatasi and P-duboscqi. Ann. Trop. Med. Parasitol. 91:267-279.

Coutinho-Abreu, I. V., and M. Ramalho-Ortigao 2010. Transmission blocking vaccines to control insect-borne diseases: a review. Mem. Inst. Oswaldo Cruz. 105:1-12.

Coutinho-Abreu, I. V., N. K. Sharma, M. Robles-Murguia, and M. Ramalho-Ortigao 2010. Targeting the midgut secreted PpChit1 reduces Leishmania major development in its natural vector, the sand fly Phlebotomus papatasi. PLoS. Negl. Trop. Dis. 4:e901.

Cunningham, M. L., R. G. Titus, S. J. Turco, and S. M. Beverley 2001. Regulation of differentiation to the infective stage of the protozoan parasite Leishmania major by tetrahydrobiopterin. Science 292:285-287.

Devenport, M., P. H. Alvarenga, L. Shao, H. Fujioka, M. L. Bianconi, P. L. Oliveira, and M. Jacobs-Lorena 2006. Identification of the Aedes aegypti peritrophic matrix protein AeIMUCI as a heme-binding protein. Biochemistry. 45:9540-9549.

Dillon, R. J., A. C. Ivens, C. Churcher, N. Holroyd, M. A. Quail, M. E. Rogers, M. B. Soares, M. F. Bonaldo, T. L. Casavant, M. J. Lehane, and P. A. Bates 2006. Analysis of ESTs from Lutzomyia longipalpis sand flies and their contribution toward understanding the insect-parasite relationship. Genomics 88:831-840.

Dillon, R. J., and R. P. Lane 1993a. Bloodmeal Digestion in the Midgut of Phlebotomus papatasi and Phlebotomus langeroni. Med. Vet. Entomol. 7:225-232.

Dillon, R. J., and R. P. Lane 1993b. Influence of Leishmania Infection on Blood-Meal Digestion in the Sandflies Phlebotomus-papatasi and P.langeroni. Parasitol. Res. 79:492-496.

Dinglasan, R. R., M. Devenport, L. Florens, J. R. Johnson, C. A. McHugh, M. Donnelly-Doman, D. J. Carucci, J. R. Yates, III, and M. Jacobs-Lorena 2009. The Anopheles gambiae adult midgut peritrophic matrix proteome. Insect Biochem. Mol. Biol. 39:125-134.

Dinglasan, R. R., and M. Jacobs-Lorena 2008. Flipping the paradigm on malaria transmission-blocking vaccines. Trends Parasitol. 24:364-370.

Dinglasan, R. R., D. E. Kalume, S. M. Kanzok, A. K. Ghosh, O. Muratova, A. Pandey, and M. Jacobs-Lorena 2007. Disruption of Plasmodium falciparum development by antibodies against a conserved mosquito midgut antigen. Proc. Natl. Acad. Sci. U S A 104:13461-13466.

Dobson, D. E., S. Kamhawi, P. Lawyer, S. J. Turco, S. M. Beverley, and D. L. Sacks 2010. Leishmania major survival in selective Phlebotomus papatasi sand fly vector requires a specific SCG-encoded lipophosphoglycan galactosylation pattern. PLoS. Pathog. 6:e1001185.

Eschenlauer, S. C., M. S. Faria, L. S. Morrison, N. Bland, F. L. Ribeiro-Gomes, G. A. Dosreis, G. H. Coombs, A. P. Lima, and J. C. Mottram 2009. Influence of parasite encoded inhibitors of serine peptidases in early infection of macrophages with Leishmania major. Cell Microbiol. 11:106-120.

Feldhaar, H., and R. Gross 2008. Immune reactions of insects on bacterial pathogens and mutualists. Microbes. Infect. 10:1082-1088.

Gossage, S. A., M. E. Rogers, and P. A. Bates 2003. Two separate growth phases during the development of Leishmania in sand flies: implications for understanding the life cycle. Int. J. Parasitol. 33:1027-1034.

GRAF, R., and H. Briegel 1989. The synthetic pathway of trypsin in the mosquito Aedes aegypti Ll (Diptera, Culicidae) and in vitro stimulation in isolated midguts. Insect. Biochem 19:129-137.

Hao, Z. R., I. Kasumba, M. J. Lehane, W. C. Gibson, J. Kwon, and S. Aksoy 2001. Tsetse immune responses and trypanosome transmission: Implications for the development of tsetse-based strategies to reduce trypanosomiasis. Proc. Natl. Acad. Sci. U S A 98:12648-12653.

Jacobs-Lorena, M., and M. M. Oo 1996. The peritrophic matrix of insects, pp. 318-332. *In* B. J. Beaty and W. C. Marquardt (eds.), The biology of disease vectors. University Press of Colorado, Colorado.

Jaramillo-Gutierrez, G., J. Rodrigues, G. Ndikuyeze, M. Povelones, A. Molina-Cruz, and C. Barillas-Mury 2009. Mosquito immune responses and compatibility between Plasmodium parasites and anopheline mosquitoes. BMC Microbiol. 9:154.:154.

Jochim, R. C., C. R. Teixeira, A. Laughinghouse, J. B. Mu, F. Oliveira, R. B. Gomes, D. E. Elnaiem, and J. G. Valenzuela 2008. The midgut transcriptome of Lutzomyia longipalpis: comparative analysis of cDNA libraries from sugar-fed, blood-fed, post-digested and Leishmania infantum chagasi-infected sand flies. BMC Genomics 9.

Joshi, P. B., B. L. Kelly, S. Kamhawi, D. L. Sacks, and W. R. Mcmaster 2002. Targeted gene deletion in Leishmania major identifies leishmanolysin (GP63) as a virulence factor. Mol. Biochem. Parasitol. 120:33-40.

Joshi, P. B., D. L. Sacks, G. Modi, and W. R. Mcmaster 1998. Targeted gene deletion of Leishmania major genes encoding developmental stage-specific leishmanolysin (GP63). Mol. Microbiol. 27:519-530.

Kamhawi, S. 2006. Phlebotomine sand flies and Leishmania parasites: friends or foes? Trends Parasitol. 22:439-445.

Kamhawi, S., G. B. Modi, P. F. P. Pimenta, E. Rowton, and D. L. Sacks 2000. The vectorial competence of Phlebotomus sergenti is specific for Leishmania tropica and is controlled by species-specific, lipophosphoglycan-mediated midgut attachment. Parasitology 121:25-33.

Kamhawi, S., M. Ramalho-Ortigao, V. M. Pham, S. Kumar, P. G. Lawyer, S. J. Turco, C. Barillas-Mury, D. L. Sacks, and J. G. Valenzuela 2004. A role for insect galectins in parasite survival. Cell 119:329-341.

Killick-Kendrick, R. 1999. The biology and control of phlebotomine sand flies. Clinics Dermatol. 17:279-289.

Killick-Kendrick, R., M. Killick-Kendrick, and Y. Tang 1994. Anthroponotic cutaneous leishmaniasis in Kabul, Afghanistan: The low susceptibility of Phlebotomus papatasi Leishmania tropica. Trans. R. Soc. Trop. Med. Hyg. 88:252-253.

Killick-Kendrick, R., M. Killick-Kendrick, Y. Tang, and P. Bastien 1996. Metacyclic promastigotes of Leishmania in the salivary glands of experimentally infected phlebotomine sandflies. Parasite. 3:55-60.

Kimblin, N., N. Peters, A. Debrabant, N. Secundino, J. Egen, P. Lawyer, M. P. Fay, S. Kamhawi, and D. Sacks 2008. Quantification of the infectious dose of Leishmania major transmitted to the skin by single sand flies. Proc. Natl. Acad. Sci. U S A 105:10125-10130.

King, D. L., and S. J. Turco 1988. A ricin agglutinin-resistant clone of Leishmania donovani deficient in lipophosphoglycan. Mol. Biochem. Parasitol. 28:285-294.

Kock, N. P., H. J. Gabius, J. Schmitz, and J. Schottelius 1997. Receptors for carbohydrate ligands including heparin on the cell surface of Leishmania and other trypanosomatids. Trop. Med. Int. Health 2:863-874.

Kumar, S., G. K. Christophides, R. Cantera, B. Charles, Y. S. Han, S. Meister, G. Dimopoulos, F. C. Kafatos, and C. Barillas-Mury 2003. The role of reactive oxygen species on Plasmodium melanotic encapsulation in Anopheles gambiae. Proc. Natl. Acad. Sci. U S A 100:14139-14144.

Kumar, S., A. Molina-Cruz, L. Gupta, J. Rodrigues, and C. Barillas-Mury 2010. A peroxidase/dual oxidase system modulates midgut epithelial immunity in Anopheles gambiae. Science. 327:1644-1648.

Lainson, R., R. D. Ward, and J. J. Shaw 1977. Leishmania in phlebotomid sandflies: VI. Importance of hindgut development in distinguishing between parasites of the Leishmania mexicana and L. braziliensis complexes. Proc. R. Soc. B 199:309-320.

Lang, T., A. Warburg, D. L. Sacks, S. L. Croft, and R. P. Lane 1991. Transmission and scanning EM-immunogold labeling of Leishmania major lipophosphoglycan in the sandfly Phlebotomus papatasi. Eur. J. Cell Biol. 55:362-372.

Lavazec, C., C. Boudin, R. Lacroix, S. Bonnet, A. Diop, S. Thiberge, B. Boisson, R. Tahar, and C. Bourgouin 2007. Carboxypeptidases B of Anopheles gambiae as targets for a Plasmodium falciparum transmission-blocking vaccine. Infect. Immun. 75:1635-1642.

Lawyer, P. G., P. M. Ngumbi, C. O. Anjili, S. O. Odongo, Y. M. Mebrathu, J. I. Githure, D. K. Koech, and C. R. Roberts 1990. development of Leishmania major in Phlebotomus duboscqi and Sergentomyia schwetzi (Diptera: Psychodidae). Am. J. Trop. Med. Hyg. 43:31-43.

Lehane, M. J. 1997. Peritrophic matrix structure and function. Annu. Rev. Entomol.

Mahoney, A. B., D. L. Sacks, E. Saraiva, G. Modi, and S. J. Turco 1999. Intra-species and stage-specific polymorphisms in lipophosphoglycan structure control Leishmania donovani-sand fly interactions. Biochemistry 38:9813-9823.

Mcconville, M. J., S. J. Turco, M. A. J. Ferguson, and D. L. Sacks 1992. Developmental Modification of Lipophosphoglycan During the Differentiation of Leishmania-Major Promastigotes to an Infectious Stage. EMBO. J. 11:3593-3600.

Meister, S., B. Agianian, F. Turlure, A. Relogio, I. Morlais, F. C. Kafatos, and G. K. Christophides 2009. Anopheles gambiae PGRPLC-mediated defense against bacteria modulates infections with malaria parasites. PLoS. Pathog. 5:e1000542.

Mukhopadhyay, N. K., K. Shome, A. K. Saha, J. R. Hassell, and R. H. Glew 1989. Heparin binds to Leishmania donovani promastigotes and inhibits protein phosphorylation. Biochem. J. 264:517-525.

Myskova, J., M. Svobodova, S. M. Beverley, and P. Volf 2007. A lipophosphoglycan-independent development of Leishmania in permissive sand flies. Microbes Infect. 9:317-324.

Oliveira, F., R. C. Jochim, J. G. Valenzuela, and S. Kamhawi 2009. Sand flies, Leishmania, and transcriptome-borne solutions. Parasitol. Int. 58:1-5.

Palanova, L., and P. Volf 1997. Carbohydrate-binding specificities and physico-chemical properties of lectins in various tissue of phlebotomine sandflies. Folia Parasitol. 44:71-76.

Peacock, C. S., K. Seeger, D. Harris, L. Murphy, J. C. Ruiz, M. A. Quail, N. Peters, E. Adlem, A. Tivey, M. Aslett, A. Kerhornou, A. Ivens, A. Fraser, M. A. Rajandream, T. Carver, H. Norbertczak, T. Chillingworth, Z. Hance, K. Jagels, S. Moule, D. Ormond, S. Rutter, R. Squares, S. Whitehead, E. Rabbinowitsch, C. Arrowsmith, B. White, S. Thurston, F. Bringaud, S. L. Baldauf, A. Faulconbridge, D. Jeffares, D. P. Depledge, S. O. Oyola, J. D. Hilley, L. O. Brito, L. R. Tosi, B. Barrell, A. K. Cruz, J. C. Mottram, D. F. Smith, and M. Berriman 2007. Comparative genomic analysis of three Leishmania species that cause diverse human disease. Nat. Genet. 39:839-847.

Pimenta, P. F. P., G. B. Modi, S. T. Pereira, M. Shahabuddin, and D. L. Sacks 1997. A novel role for the peritrophic matrix in protecting Leishmania from the hydrolytic activities of the sand fly midgut. Parasitology 115:359-369.

Pimenta, P. F. P., E. MB. Saraiva, E. Rowton, G. B. Modi, L. A. Garraway, S. M. Beverley, S. J. Turco, and D. L. Sacks 1994. Evidence that the vectorial competence of phlebotomine sand flies for different species of Leishmania is controlled by structural polymorphisms in the surface lipophosphoglycan. Proc. Natl. Acad. Sci. USA. 91:9155-9159.

Pimenta, P. F. P., S. J. Turco, M. J. Mcconville, P. G. Lawyer, P. V. Perkins, and D. L. Sacks 1992. Stage-Specific Adhesion of Leishmania Promastigotes to the Sandfly Midgut. Science 256:1812-1815.

Pinto-da-Silva, L. H., P. Fampa, D. C. Soares, S. M. P. Oliveira, T. Souto-Padron, and E. M. Saraiva 2005. The 3A1-La monoclonal antibody reveals key features of Leishmania (L) amazonensis metacyclic promastigotes and inhibits procyclics attachment to the sand fly midgut. Int. J. Parasitol. 35:757-764.

Pitaluga, A. N., V. Beteille, A. R. Lobo, J. R. Ortigao-Farias, A. M. Davila, A. A. Souza, J. M. Ramalho-Ortigao, and Y. M. Traub-Cseko 2009. EST sequencing of blood-fed and Leishmania-infected midgut of Lutzomyia longipalpis, the principal visceral leishmaniasis vector in the Americas. Mol. Genet. Genomics. 282:307-317.

Rabinovich, G. A., and A. Gruppi 2005. Galectins as immunoregulators during infectious processes: from microbial invasion to the resolution of the disease. Parasite Immunology 27:103-114.

Ramalho-Ortigao, J. M., S. Kamhawi, M. B. Joshi, D. Reynoso, P. G. Lawyer, D. M. Dwyer, D. L. Sacks, and J. G. Valenzuela 2005. Characterization of a blood activated chitinolytic system in the midgut of the sand fly vectors Lutzomyia longipalpis and Phlebotomus papatasi. Insect Mol. Biol. 14:703-712.

Ramalho-Ortigao, J. M., S. Kamhawi, E. D. Rowton, J. M. C. Ribeiro, and J. G. Valenzuela 2003. Cloning and characterization of trypsin- and chymotrypsin-like proteases from the midgut of the sand fly vector Phlebotomus papatasi. Insect Biochem. Mol. Biol. 33:163-171.

Ramalho-Ortigao, J. M., A. N. Pitaluga, E. L. Telleria, C. Marques, A. A. Souza, and Y. M. Traub-Cseko 2007a. Cloning and characterization of a V-ATPase subunit C from the American visceral leishmaniasis vector Lutzomyia longipalpis modulated during development and blood ingestion. Mem. Inst. Oswaldo Cruz 102:509-515.

Ramalho-Ortigao, J. M., and Y. M. Traub-Cseko 2003. Molecular characterization of Llchit1, a midgut chitinase cDNA from the leishmaniasis vector Lutzomyia longipalpis. Insect Biochem. Mol. Biol. 33:279-287.

Ramalho-Ortigao, M., R. C. Jochim, J. M. Anderson, P. G. Lawyer, V. M. Pham, S. Kamhawi, and J. G. Valenzuela 2007b. Exploring the midgut transcriptome of Phlebotomus papatasi: comparative analysis of expression profiles of sugar-fed, blood-fed and Leishmania major-infected sandflies. BMC Genomics 8.

Ramirez, J. L., L. S. Garver, and G. Dimopoulos 2009. Challenges and approaches for mosquito targeted malaria control. Curr. Mol. Med. 9:116-130.

Rogers, M. E., and P. A. Bates 2007. Leishmania manipulation of sand fly feeding behavior results in enhanced transmission. Plos Pathog. 3:818-825.

Rogers, M. E., M. L. Chance, and P. A. Bates 2002. The role of promastigote secretory gel in the origin and transmission of the infective stage of Leishmania mexicana by the sandfly Lutzomyia longipalpis. Parasitology 124:495-507.

Rogers, M. E., M. Hajmova, M. B. Joshi, J. Sadlova, D. M. Dwyer, P. Volf, and P. A. Bates 2008. Leishmania chitinase facilitates colonization of sand fly vectors and enhances transmission to mice. Cell. Microbiol. 10:1363-1372.

Rogers, M. E., T. Ilg, A. V. Nikolaev, M. A. J. Ferguson, and P. A. Bates 2004. Transmission of cutaneous leishmaniasis by sand flies is enhanced by regurgitation of fPPG. Nature 430:463-467.

Sacks, D., and S. Kamhawi 2001. Molecular aspects of parasite-vector and vector-host interactions in Leishmaniasis. Ann. Rev. Microbiol. 55:453-483.

Sacks, D. L. 2001. Leishmania-sand fly interactions controlling species-specific vector competence. Cell. Microbiol. 3:189-196.

Sacks, D. L., T. Brodin, and S. J. Turco 1990. Developmental modification of the lipophosphoglycan from Leishmania major promastigotes during metacyclogenesis. Mol. Biochem. Parasitol. 42:225-234.

Sacks, D. L., G. Modi, E. Rowton, G. Spath, L. Epstein, S. J. Turco, and S. M. Beverley 2000. The role of phosphoglycans in Leishmania-sand fly interactions. Proc. Natl. Acad. Sci. USA 97:406-411.

Sacks, D. L., P. F. P. Pimenta, M. J. Mcconville, P. Schneider, and S. J. Turco 1995. Stage-specific binding of Leishmania donovani to the sand fly vector midgut is regulated by conformational changes in the abundant surface lipophosphoglycan. J. Exp. Med. 181:685-697.

Sadlova, J. 1999. The life history of Leishmania (kinetoplastida: Trypanosomatidae). Acta Soc. Zool. Bohem. 63:331-366.

Sadlova, J., and P. Volf 2009. Peritrophic matrix of Phlebotomus duboscqi and its kinetics during Leishmania major development. Cell Tissue Res. 337:313-325.

Sadlova, J., and P. Volf 1999. Occurrence of Leishmania major in sandfly urine. Parasitology 118:455-460.

Sant'Anna, M. R., H. az-Albiter, M. Mubaraki, R. J. Dillon, and P. A. Bates 2009. Inhibition of trypsin expression in Lutzomyia longipalpis using RNAi enhances the survival of Leishmania. Parasit. Vectors. 2:62.

Saraiva, E. MB., P. F. P. Pimenta, T. N. Brodin, E. Rowton, G. B. Modi, and D. L. Sacks 1995. Changes in lipophosphoglycan and gene expression associated with the development of Leishmania major in Phlebotomus papatasi. Parasitology. 111:275-287.

Schlein, Y., and R. L. Jacobson 1998. Resistance of Phlebotomus papatasi to infection with Leishmania donovani is modulated by components of the infective bloodmeal. Parasitology 117:467-473.

Schlein, Y., R. L. Jacobson, and G. Messer 1992. Leishmania Infections Damage the Feeding Mechanism of the Sandfly Vector and Implement Parasite Transmission by Bite. Proc. Natl. Acad. Sci. USA 89:9944-9948.

Schlein, Y., R. L. Jacobson, and J. Shlomai 1991. Chitinase secreted by Leishmania functions in the sandfly vector. Proc. R. Soc. Lond. B 245:121-126.

Schlein, Y., and H. Romano 1986. Leishmania major and L.donovani: effects on proteolytic enzymes of Phlebotomus papatasi (Diptera: Psychodidae). Exp. Parasitol. 62:376-380.

Secundino, N., N. Kimblin, N. C. Peters, P. Lawyer, A. A. Capul, S. M. Beverley, S. J. Turco, and D. Sacks 2010. Proteophosphoglycan confers resistance of Leishmania major to midgut digestive enzymes induced by blood feeding in vector sand flies. Cell. Microbiol. 12:906-918.

Secundino, N. F. C., I. Eger-Mangrich, E. M. Braga, M. M. Santoro, and P. F. P. Pimenta 2005. Lutzomyia longipalpis peritrophic matrix: Formation, structure, and. chemical composition. J. Med. Entomol. 42:928-938.

Shao, L., M. Devenport, H. Fujioka, A. Ghosh, and M. Jacobs-Lorena 2005. Identification and characterization of a novel peritrophic matrix protein, Ae-Aper50, and the microvillar membrane protein, AEG12, from the mosquito, Aedes aegypti. Insect Biochem. Mol. Biol. 35:947-959.

Shatova, S. M., M. A. Shulga, V. M. Safjanova, and A. A. Avakjan 1984. Comparative electron microscope study of Leishmania major and L. tropica at the experimental infection of the sandfly Phlebotomus papatasi. Parazitologija 18:154-159.

Soares, R. P. P., M. E. Macedo, C. Ropert, N. F. Gontijo, I. C. Almeida, R. T. Gazzinelli, P. F. P. Pimenta, and S. J. Turco 2002. Leishmania chagasi: lipophosphoglycan characterization and binding to the midgut of the sand fly vector Lutzomyia longipalpis. Mol. Biochem. Parasitol. 121:213-224.

Spath, G. F., L. F. Lye, H. Segawa, D. L. Sacks, S. J. Turco, and S. M. Beverley 2003. Persistence without pathology in phosphoglycan-deficient Leishmania major. Science 301:1241-1243.

Svobodova, M., P. A. Bates, and P. Volf 1997. Detection of lectin activity in Leishmania promastigotes and amastigotes. Acta Trop. 68:23-35.

Telleria, E. L., A. P. de Araujo, N. F. Secundino, C. M. vila-Levy, and Y. M. Traub-Cseko 2010. Trypsin-Like Serine Proteases in Lutzomyia longipalpis - Expression, Activity and Possible Modulation by Leishmania infantum chagasi. PLoS. One. 5:e10697.

Telleria, E. L., A. N. Pitaluga, J. R. Ortigao-Farias, A. P. O. de Araujo, J. M. Ramalho-Ortigao, and Y. M. Traub-Cseko 2007. Constitutive and blood meal-induced trypsin genes in Lutzomyia longipalpis. Arch. Insect Biochem. Physiol. 66:53-63.

Thomas, J. R., M. J. Mcconville, J. E. Thomasoates, S. W. Homans, M. A. J. Ferguson, P. AJ. Gorin, K. D. Greis, and S. J. Turco 1992. Refined Structure of the Lipophosphoglycan of Leishmania- Donovani. J. Biol. Chem. 267:6829-6833.

Turco, S. J., and A. Descoteaux 1992. The lipophosphoglycan of Leishmania parasites. Annu. Rev. Microbiol. 46:65-94.

Volf, P., M. Hajmova, J. Sadlova, and J. Votypka 2004. Blocked stomodeal valve of the insect vector: similar mechanism of transmission in two trypanosomatid models. Int. J. Parasitol. 34:1221-1227.

Volf, P., and J. Myskova 2007. Sand flies and *Leishmania*: specific versus permissive vectors. Trends Parasitol. 23:91-92.

Volf, P., L. Palanova, and M. Svobodova 1995. Midgut lectins of sandflies. Bol. Dir. Malariol. y San. Amb. 35:371-374.

Volf, P., M. Svobodova, and E. Dvorakova 2001. Bloodmeal digestion and Leishmania major infections in Phlebotomus duboscqi: effect of carbohydrates inhibiting midgut lectin activity. Med. Vet. Entomol. 15:281-286.

Wallbanks, K. R., G. A. Ingram, and D. H. Molyneux 1986. The agglutination of erythrocytes and Leishmania parasites by sandfly gut extracts: evidence for lectin activity. Trop. Med. Parasitol. 37:409-413.

Walters, L. I., K. P. Irons, G. B. Modi, and R. B. Tesh 1992. Refractory barriers in the sandfly Phlebotomus papatasi (Diptera: Psychodidae) to infection with Leishmania panamensis. Am. J. Trop. Med. Hyg. 46:211-228.

Warburg, A., and Y. Schlein 1986. The effect of postbloodmeal nutrition of Phlebotomus papatasi on the transmission of Leishmania major. Am. J. Trop. Med. Hyg. 35:926-930.

Warburg, A., R. B. Tesh, and D. McMahon-Pratt 1989. Studies on the attachemnent of Leishmania flagella to sand fly midgut epithelium. J. Protozool. 36:613-617.

Warr, E., S. Das, Y. Dong, and G. Dimopoulos 2008. The Gram-negative bacteria-binding protein gene family: its role in the innate immune system of anopheles gambiae and in anti-Plasmodium defence. Insect Mol. Biol. 17:39-51.

OPEN ACCESS Freely available online

PLoS NEGLECTED TROPICAL DISEASES

Stage-Specific Adhesion of *Leishmania* Promastigotes to Sand Fly Midguts Assessed Using an Improved Comparative Binding Assay

Raymond Wilson[1], Michelle D. Bates[1], Anna Dostalova[2], Lucie Jecna[2], Rod J. Dillon[1], Petr Volf[2], Paul A. Bates[3]*

1 Liverpool School of Tropical Medicine, Liverpool, United Kingdom, **2** Department of Parasitology, Charles University, Prague, Czech Republic, **3** School of Health and Medicine, Lancaster University, Lancaster, United Kingdom

Abstract

Background: The binding of *Leishmania* promastigotes to the midgut epithelium is regarded as an essential part of the life-cycle in the sand fly vector, enabling the parasites to persist beyond the initial blood meal phase and establish the infection. However, the precise nature of the promastigote stage(s) that mediate binding is not fully understood.

Methodology/Principal Findings: To address this issue we have developed an *in vitro* gut binding assay in which two promastigote populations are labelled with different fluorescent dyes and compete for binding to dissected sand fly midguts. Binding of procyclic, nectomonad, leptomonad and metacyclic promastigotes of *Leishmania infantum* and *L. mexicana* to the midguts of blood-fed, female *Lutzomyia longipalpis* was investigated. The results show that procyclic and metacyclic promastigotes do not bind to the midgut epithelium in significant numbers, whereas nectomonad and leptomonad promastigotes both bind strongly and in similar numbers. The assay was then used to compare the binding of a range of different parasite species (*L. infantum*, *L. mexicana*, *L. braziliensis*, *L. major*, *L. tropica*) to guts dissected from various sand flies (*Lu. longipalpis*, *Phlebotomus papatasi*, *P. sergenti*). The results of these comparisons were in many cases in line with expectations, the natural parasite binding most effectively to its natural vector, and no examples were found where a parasite was unable to bind to its natural vector. However, there were interesting exceptions: *L. major* and *L. tropica* being able to bind to *Lu. longipalpis* better than *L. infantum*; *L. braziliensis* was able to bind to *P. papatasi* as well as *L. major*; and significant binding of *L. major* to *P. sergenti* and *L. tropica* to *P. papatasi* was observed.

Conclusions/Significance: The results demonstrate that *Leishmania* gut binding is strictly stage-dependent, is a property of those forms found in the middle phase of development (nectomonad and leptomonad forms), but is absent in the early blood meal and final stages (procyclic and metacyclic forms). Further they show that although gut binding may be necessary for parasite establishment, in several vector-parasite pairs the specificity of such *in vitro* binding alone is insufficient to explain overall vector specificity. Other significant barriers to development must exist in certain refractory *Leishmania* parasite-sand fly vector combinations. A re-appraisal of the specificity of the *Leishmania*-sand fly relationship is required.

Citation: Wilson R, Bates MD, Dostalova A, Jecna L, Dillon RJ, et al. (2010) Stage-Specific Adhesion of *Leishmania* Promastigotes to Sand Fly Midguts Assessed Using an Improved Comparative Binding Assay. PLoS Negl Trop Dis 4(9): e816. doi:10.1371/journal.pntd.0000816

Editor: Genevieve Milon, Institut Pasteur, France

Received February 22, 2010; **Accepted** August 10, 2010; **Published** September 7, 2010

Funding: This work was supported by Project Grant 078937 from The Wellcome Trust (http://www.wellcome.ac.uk/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: p.bates@lancaster.ac.uk

Introduction

Medically important protozoan parasites of the genus *Leishmania* have a two host life-cycle, alternating between various mammalian hosts and female haematophagous phlebotomine sand flies. The life-cycle within the sand fly gut is complex, varies between subgenera *Leishmania* and *Viannia*, and includes a number of distinct morphological forms given various names [1–5]. In the subgenus *Leishmania*, amastigote forms are released from mammalian macrophages in the midgut of the fly after ingestion of a blood meal, and transform into proliferative procyclic promastigote forms. Around 48–72 hours later, these develop into non-dividing and long nectomonad promastigotes, which escape from the peritrophic matrix-encased blood meal into the midgut lumen. Here they develop into leptomonad promastigotes (a synonym for short nectomonads [2,3]), which enter another proliferative cycle, and by day 4 post-blood meal comprise the majority of the parasite population. The final phase is the transformation of leptomonad forms into mammalian infective stages, metacyclic promastigotes, and by 7–10 days post-blood feeding over 60% of the parasite population in the region of the stomodeal valve (junction between midgut and foregut) is comprised of this form [4,5].

One essential event in the establishment of *Leishmania* infections in the sand fly is the attachment of the parasites to the midgut epithelium. By anchoring themselves to the midgut the parasites help to prevent their expulsion from the gut during defecation, and

Author Summary

Many infectious diseases such as leishmaniasis are transmitted to people by biting insects, in this case by female sand flies. To control this and similar diseases we need to understand why particular species of sand fly transmit particular species of *Leishmania*. One important feature of the *Leishmania* parasite-sand fly interaction is the ability of the parasite to bind to the midgut wall of the fly, as it is within the gut that the parasite lives. Here we have studied the specificity of this interaction and report two main findings. The first is that only specific stages in the parasite life-cycle are capable of binding to the gut. The second is that, providing these life-cycle stages are analysed, parasite species that can be transmitted by particular sand flies are always capable of binding to their guts, but in some cases they are also capable of binding to non-transmitting sand fly species. This shows that gut binding by parasites is necessary but not sufficient to explain transmission. This research advances our understanding of *Leishmania* biology, but also shows us that there are further aspects that need to be investigated before we can fully understand the *Leishmania*-sand fly relationship.

it has been postulated that this binding is the main determinant of parasite-vector specificity [6–9]. The mechanism of binding has been most intensively studied in *L. major* infections of *Phlebotomus papatasi* [6,7,10–12]. The promastigote surface glycoconjugate lipophosphoglycan (LPG) was implicated as the parasite molecule involved in midgut attachment of *L. major* [6]. In this parasite-vector combination, competitive inhibition with oligosaccharides derived from LPG [6], anti-LPG antibodies [6], and mutant strains that either lack LPG entirely [9] or produce LPG that lacks the carbohydrate side-chain moieties present on wild-type *L. major* LPG [12], all caused a significant reduction in the binding of *L. major* to *P. papatasi* midguts in vitro and/or a reduction/ablation in the parasite's ability to complete its life-cycle in the sand fly. A corresponding galactose-binding lectin or galectin (PpGalec) has been described on the midgut of *P. papatasi* that binds to *L. major* LPG, and antibodies against PpGalec inhibit promastigote attachment [13]. Similar, although less complete, analyses have been performed on various additional vector-parasite combinations, leading to the hypothesis that LPG is the key molecule for midgut attachment in all *Leishmania* species [14]. One unexplained observation is that attachment usually occurs via the flagellum, which is seen to insert between the microvilli, whereas LPG is found over the whole surface of promastigotes. The reasons may be mechanical, the flagellum being at the anterior of the cell and more easily able to fit in between the microvilli, but could indicate that other molecules have a role to play in attachment. However, in subsequent work it was shown that LPG is not essential for completion of the parasite life-cycle in all *Leishmania*/sand fly infections [15,16]. This suggests either that attachment is not always obligatory for life-cycle progression, in contradiction to the current paradigm, or that an alternative or supplementary non-LPG mediated attachment mechanism exists.

Non-LPG mediated attachment of promastigotes to the sand fly gut has only been observed in sand flies termed "permissive vectors" [15]. This concept relates to the subdivision of vector sand fly species into two broad categories, permissive and specific/restricted/refractory vectors [17–19]. Specific vectors only allow the maturation and transmission of a single *Leishmania* species. For example, *P. papatasi* supports development of *L. major* only, and is refractory to all other species tested [7]. Similarly, *P. sergenti* supports the maturation of typical *L. tropica* strains [20,21]. The majority of sand flies, however, fall into the permissive category, and multiple *Leishmania* species are able to survive and mature within the gut of such a species if given the chance by experiment or nature. Examples of permissive flies are the New World species, *Lutzomyia longipalpis* [4,22,23], and the Old World species *P. argentipes*, *P. arabicus*, *P. halepensis* and *P. perniciosus* [7,15,16,21,24], which allow the maturation of practically all *Leishmania* species tested under experimental conditions. For example, the Old World parasite *L. major* is able to complete its development in the New World vector *Lu. longipalpis* [22], which is the natural vector for *L. infantum* (syn. *chagasi*). The current evidence indicates that infections of specific vectors have a strict dependence on species-specific LPG for binding the parasites to the midgut. However, in permissive vectors the picture is less clear, LPG does not appear to be essential, and by implication there appears to be an additional LPG-independent binding mechanism in these vectors [15]. In this study, we examined midgut-binding in both specific and permissive vectors.

Materials and Methods

Parasites and sand flies

The following parasite isolates were used in this study: *Leishmania braziliensis* MHOM/BR/84/LTB300; *L. infantum* (syn. *chagasi*) MHOM/BR/76/M4192; *L. major* LV561 MHOM/IL/67/LRC-L137; *L. mexicana* MNYC/BZ/62/M379; and *L. tropica* SU23 MHOM/TR/98/HM. Laboratory colonies of three sand fly species were used: *Lutzomyia longipalpis* (origin from Jacobina, Brazil); *Phlebotomus papatasi* (Turkey); and *P. sergenti* (Turkey).

Generation of *Leishmania* life-cycle stages

Leishmania parasites were assigned to life-cycles stages according to the classification of Rogers *et al.* [4]. *L. mexicana* lesion amastigotes were transferred into promastigote culture medium (M199 medium containing 25 μg/ml gentamicin sulphate (Sigma), 1x BME vitamins (from 100x stock, Gibco), 20% foetal bovine serum (FBS; Sigma), and 2% urine) at 26°C and promastigote cultures enriched for procyclic, nectomonad, leptomonad and metacyclic forms were collected after 1, 2, 4 and 8 days of in vitro culture, respectively [4]. For *L. infantum* it was not possible to obtain sufficient quantity of promastigotes direct from mouse spleen homogenate amastigotes. For this species, amastigotes of *L. infantum* were transferred into promastigote culture medium and counted daily until exponential growth of promastigotes was observed (2–3 days). The parasites were then passaged at a concentration of 5×10^5/ml into fresh medium and cultured for 2–3 days or 7 days to produce leptomonad-enriched or metacyclic-enriched cultures, respectively. Procyclic and nectomonad promastigotes were derived from metacyclic cultures that had been centrifuged and resuspended in Grace's Insect Medium (GIM) containing 20% FBS, pH 5.5 and cultured at 32°C until transformation to amastigote forms had occurred (1–2 days). The resulting amastigotes were then transferred back into promastigote culture medium and cultured at 26°C for 24 hours to produce procyclic-enriched cultures or 48 hours to produce nectomonad-enriched cultures. For additional species of *Leishmania* a methodology similar to that used with *L. infantum* was employed. For species-species comparisons, leptomonad-enriched cultures were prepared for each parasite isolate.

Parasites from the enriched cultures were cryopreserved in 10% dimethylsulphoxide, 1 ml aliquots stored under liquid nitrogen, and thawed as required for binding assays. Thawed aliquots were washed twice in promastigote culture medium to remove

dimethylsulphoxide prior to use. Viability was then checked microscopically and very few if any dead (non-motile) promastigotes were observed in the majority of cases. However, in rare instances where more than 5% of the promastigotes were not viable such parasites were not used in experiments. This precaution was taken in case the original population was in poor condition, rather than due to the presence of dead parasites per se, as these did not bind to midguts under the conditions of the assay. All of the populations used were enriched to at least 80% of the relevant life-cycle stage, with the exception of *L. infantum* nectomonad promastigotes, which could only be obtained to 60% purity. All procedures involving animals were performed in accordance with UK Government (Home Office) and EC regulations, and were approved by the University of Liverpool Animal Welfare Committee.

Midgut-binding assay

Lu. longipalpis and *P. sergenti* midguts were dissected in GIM four days after blood-feeding. *P. papatasi* were dissected 6 days after blood-feeding due to the blood meal digestion period in this species being slightly longer. Hindgut, foregut and Malpighian tubules were removed and the midgut carefully opened longitudinally using a bevelled, fine glass needle. The needles were made from borosilicate 3·5 inch hematocrit capillary tubes (Drummond, USA) drawn into fine tips using a PC-10 Narishige Puller. Before use the needles were bevelled using an EG-44 Narishige microgrinder. Parasites were labelled for 1 hr at 26°C with either 0·04% (v/v) Syto21 Green or 0·2% Syto40 Blue fluorescent dyes (Molecular Probes), washed 3 times in M199 medium and twice in GIM, then resuspended in GIM at a concentration of 1×10^8 cells/ml. The two parasite populations were mixed in equal quantities before overlaying 10 µl of the mixture on each gut and incubating at 26°C for 45 minutes in a humid chamber. The gut was washed 5 times by transferring to fresh drops of GIM, placed in 2 µl of cooled CyGel (Biostatus Ltd.) on a microscope slide, and carefully flattened so that no areas of the gut were folded over. The slide was warmed to solidify the CyGel, thereby fixing the gut in place, before mounting in Prolong Gold antifade reagent. Bound parasites of both populations were visualised simultaneously using an Olympus BX60 fluorescence microscope at x400 magnification with a CFP/YFP-A filter set (Semrock Inc.), the number of parasites from each population counted in one field of view per midgut, and their relative proportions determined. No consistent variation in binding along the length of the dissected guts was observed, but damaged areas were not used. Smears were made from the labelled parasites, stained with Giemsa's stain and measured to determine the proportion of each life-cycle stage in each culture. Each experiment was repeated with the colour labelling of the parasite populations switched, so that any potential dye-specific bias was negated (technical repeat), and each pair of comparisons was performed at least twice (biological repeat). The difference between the proportions of bound parasites of two populations was evaluated statistically by means of an ANOVA test using the R software (http://www.r-project.org/).

Results and Discussion

Development of the parasite-binding assay

Previous investigations of *Leishmania*-sand fly interactions have shown that, from around the time of escape from the peritrophic matrix and continuing throughout the course of the infection, a proportion of promastigote forms bind to the gut epithelium by intercalating their flagella between the microvilli and/or direct contact of the cell body [2,25]. In this study a comparative binding assay was developed and utilised to investigate the stage- and species-specificity of gut binding. This assay differs from previous assays in several respects. First, it utilises direct microscopic examination of a mixture of two differentially stained parasite populations to individual dissected sand fly midguts (Fig. 1). This relative binding methodology was used because, even under optimum conditions, individual sand flies (and their midguts) will be in slightly different physiological states, of slightly different ages, and dissection of a fully intact gut (required for reliable absolute numbers) is also technically demanding. Considerable individual

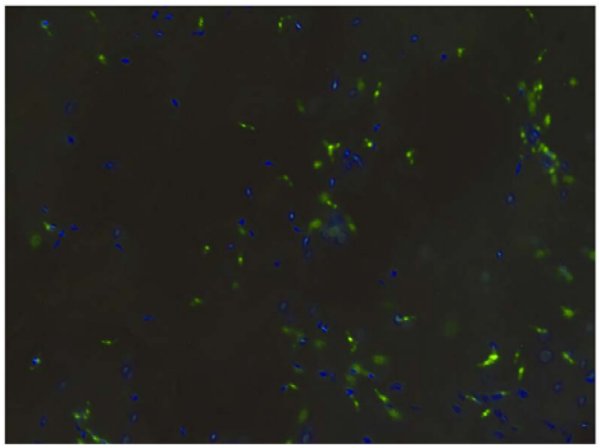


Figure 1. Binding of two *Leishmania* populations to a sand fly midgut. In the example shown, *L. mexicana* nectomonad promastigotes are labelled with Syto Blue and *L. mexicana* leptomonads with Syto Green, and are bound to a flattened *Lu. longipalpis* midgut. The gut was photographed after washing, and all the bound promastigotes were alive as revealed by movements of their cell bodies. Some areas with fewer promastigotes can be seen, but these had no obvious distribution.
doi:10.1371/journal.pntd.0000816.g001

variation in the numbers of parasites bound to sand fly guts has been noted before [7,20]. In preliminary experiments the absolute numbers of single (unmixed) populations of promastigotes binding to midguts were determined, and as expected large variations in binding to the surface midgut epithelium were noted, for example ranging from 91 to 847 per field of view (average $=433\pm44{\cdot}2$ SEM) for *L. infantum* with *Lu. longipalpis*. It is also relevant to note previous observations that the binding of *L. major* to *P. papatasi* midguts was not spatially uniform, occurring on galectin-expressing patches of cells of varying size [13]. The protocol developed in the current study minimises the biological variability of previous studies associated with comparison of binding of different populations to different midguts.

Another feature of these experiments was that they were performed using guts dissected from flies at an appropriate time post-blood feeding with respect to *Leishmania* development. Surprisingly, all previous studies have used "sugar-fed" flies, i.e. flies that had emerged after pupation but had not yet taken a blood meal. However, blood feeding is known to exert profound physiological changes on the midgut epithelium, the secretion of digestive enzymes and production of the peritrophic matrix being two obvious examples. Thus, it is unwise to assume that binding of parasites to sugar-fed flies would accurately mimic binding in an infected sand fly following blood feeding, the analogous situation under natural conditions. This is further supported by possible upregulation of galectin expression in *P. papatasi* following blood feeding [13]. Therefore, the results in this study were all derived from assays performed with guts dissected from flies 4 (the majority) or 6 (*P. papatasi*) days after blood feeding, these being the times at which maximum binding of promastigotes has been reported to occur *in vivo*.

Finally, previous *in vitro* midgut binding assays have all been performed in phosphate buffered saline [6,7,20,26,27). In addition to being non-physiological, and as has also been noted by Kamhawi *et al.* [20], we found sand fly midguts were not very stable in this solution and their structure began to deteriorate after 10 minutes incubation at 26°C. Therefore, we performed the assays in GIM, which as a customised insect culture medium resolved this problem, with the guts maintaining their structural integrity and remaining amenable to manipulation for at least 4 hours. Moreover, GIM is also a suitable medium for the growth and maintenance of *Leishmania* promastigotes [28].

The assay described here has a number of advantages over the established *in vitro* assay [6,29]. One of the major differences is in the method of quantification. Here, we determined the number of *Leishmania* parasites from two distinct populations that were bound to the same midgut. By expressing these numbers as a ratio of the numbers of each parasite population bound per gut, a large amount of the variation between samples was eliminated. Individual variation is reduced in the new assay by only counting parasites from one microscope field of view per gut instead of homogenising the whole gut and counting the released parasites on a haemocytometer, which, due to differences in sizes of the guts and quality of dissections, introduces a source of variation. In contrast to previously described *in vitro* assays, this one is suitable also for studies on midguts from wild-caught sand flies. Nevertheless, we strongly encourage researchers to use colonized females as they represent more standardized material.

Life-cycle stages of *L. infantum* and *L. mexicana* involved in binding to the midgut

Leishmania promastigotes occur in four major developmental forms in sand flies [4]: procyclic, nectomonad, leptomonad and metacyclic promastigotes. In addition there is a small population of haptomonad promastigotes and paramastigotes attached to the cuticular surfaces of the stomodeal valve and foregut by specialised hemidesmosomal structures, a completely different attachment mechanism to that being examined in this study, but these are not involved in attachment to the midgut. To determine the life-cycle stages involved in binding to the sand fly midgut, we initially attempted to measure parasites bound *in situ* to midguts of *Lu. longipalpis*. These experiments indicated that nectomonad and leptomonad promastigotes were the major forms responsible for binding to the midgut (data not shown). However, a significant technical problem in measuring such bound parasites *in situ* is that accurate measurement relies on the parasite being perfectly flat in relation to the plane of focus. Any deviation in the vertical angle of the parasite results in a shortening of the perceived length. Therefore, we adopted the comparative binding assay with *Lu. longipalpis* to determine the binding of the different developmental forms. Promastigote populations of *L. mexicana* and *L. infantum* enriched for the four different types of promastigotes were prepared and used. Comparing the binding of a population enriched for nectomonad promastigotes with a leptomonad-enriched population showed that nectomonads bound to the midgut in higher numbers in both *L. mexicana* and *L. infantum* (Fig. 2), although the difference was not statistically significant in *L. mexicana*. The leptomonad-enriched populations bound at levels $52{\cdot}0\pm9{\cdot}9\%$ ($P<0{\cdot}01$) and $78{\cdot}7\pm16{\cdot}3\%$ ($P=0{\cdot}97$) of the nectomonad-enriched populations in *L. infantum* and *L. mexicana*, respectively. The nectomonad-enriched cultures were, therefore, used as the standard to which the other life-cycle stages were compared. In confirmation of the observations made with the direct measurements of bound parasites, there was little significant binding of either procyclic or metacyclic promastigote-enriched populations to the midgut in comparison to the nectomonad population in either *L. infantum* (procyclic $7{\cdot}1\pm1{\cdot}5\%$, $P<0{\cdot}01$; metacyclic $6{\cdot}5\pm3{\cdot}7\%$, $P<0{\cdot}01$) or *L. mexicana* (procyclic $0{\cdot}55\pm0{\cdot}55\%$, $P<0{\cdot}01$; metacyclic $5{\cdot}6\pm4{\cdot}6\%$, $P<0{\cdot}01$). Similarly, there was minimal binding of metacyclic promastigotes relative to leptomonad promastigotes. Further, direct observation indicated that virtually all of the bound parasites from the procyclic and metacyclic populations were derived from contaminating nectomonad and leptomonad promastigotes, respectively, that made up to 20% of the respective starting populations. Total numbers of bound parasites per field of view ranged from less than 5 to a few hundred depending on which combination of cell types were being examined. Thus numbers of bound procyclic and metacyclic promastigotes were on average 1–2 parasites per field, whereas averages for nectomonad and leptomonad promastigotes ranged from 50 to 200 parasites per field. Since the assay used is a comparative binding assay, it was possible the procyclic and metacyclic forms were able to bind but were being out-competed by the strongly binding nectomonad and leptomonad forms. This was, however, found not to be the case in experiments comparing the binding of procyclic versus metacyclic populations (data not shown), where neither were observed binding to the guts in significant numbers.

The results presented here provide a detailed analysis of the life-cycle stages involved in midgut-binding, with the only two forms observed binding to the midgut epithelium being the nectomonad and leptomonad promastigotes. Nectomonads are the predominant form 2–4 days post-blood feeding, while the leptomonad forms are present in high numbers from 4 days onwards [4,5]. The main purpose of binding to the sand fly midguts is thought to be in preventing expulsion from the gut due to peristalsis, particularly during defecation of the blood meal [14], therefore, the binding data corresponds well with the promastigote forms present at the

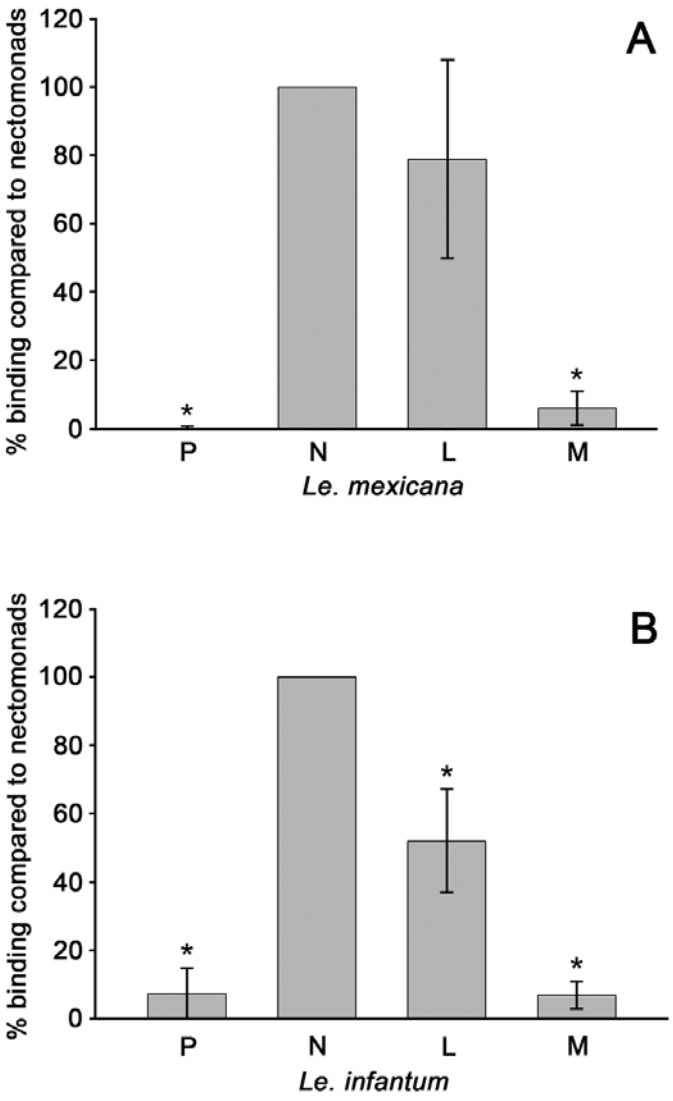


Figure 2. Comparative binding of enriched parasite populations to *Lu. longipalpis* midguts. All populations were compared to the binding of the nectomonad-enriched culture. (A) *L. mexicana* (11 experiments) (B) *L. infantum* (5 experiments). Cultures were enriched for: procyclic promastigotes (P); nectomonad promastigotes (N); leptomonad promastigotes (L) and metacyclic promastigotes (M). Each bar represents the mean percentage ± S.D. and the asterisks indicate a significant difference to the nectomonad population (>20 dissected guts examined per experiment).
doi:10.1371/journal.pntd.0000816.g002

time binding to the gut is most expected. The predominant form bound to the sand fly gut at a particular point in development will be the product of various factors. From these results we can be confident that nectomonad promastigotes will constitute the first population of bound promastigotes, since these preceed leptomonad forms. However, since the detachment rate is unknown the subsequent population is difficult to predict. Attachment could be effectively permanent if the detachment rate is very low, or it could be very dynamic with promastigotes continuously attaching and detaching. In the former scenario replacement of the nectomonad population might be relatively slow even though the total gut population becomes dominated by leptomonad forms; in the latter scenario the replacement of nectomonad with leptomonad forms could occur relatively quickly. Resolution of this will likely require

dissection or real time imaging of infected flies combined with stage specific markers for nectomonad and leptomonad forms, which is not currently possible but an interesting area for future investigation.

The inability of metacyclic promastigotes to bind is intuitively expected, since these are the infective forms of *Leishmania* and a propensity for strong binding to the gut lining would be a considerable disadvantage at this stage of its life-cycle. The loss in binding capacity of metacyclics has been attributed to changes in the structure of the major parasite surface molecule, LPG, during metacyclogenesis [6,27,29–31]. The LPG on the surface of metacyclics has a much longer phosphoglycan backbone and changes the nature of its terminally exposed sugar residues during metacyclogenesis. Different *Leishmania* species have different LPG modification mechanisms: *L. major* caps the terminal galactose residues with arabinose [6]; *L. donovani* and *L. infantum* exhibit conformational changes that mask the terminal sugars [26,29]; and *L. braziliensis* adds glucose side-chains [27]. Another possible factor influencing binding in vivo is the secretion of phosphoglycans. Although LPG itself is not shed by promastigotes in vivo, filamentous proteophosphoglycan(s) are [32], and these could act as competitive inhibitors of LPG-mediated binding since they share similar phosphoglycans. This could assist in the detachment of nectomonad or leptomonad promastigotes and/or prevent the re-attachment of metacyclic promastigotes. However, changes in LPG/PPG structure appear to be insufficient by themselves to explain lack of metacyclic binding in permissive vectors where non-LPG mediated attachment mechanisms have been described.

Procyclic promastigotes, on the other hand, may be unable to bind to the midgut due to a lack of (sufficient) LPG on their surface, since *in situ* immunohistological studies in *L. major* failed to detect any surface LPG until day 3 post-blood meal [33], a time point which corresponds with the appearance of the nectomonad promastigotes. There is also the possibility of procyclic promastigotes being unable to bind to the midgut for mechanical reasons. Morphologically, procyclic promastigotes differ from nectomonad and leptomonad promastigotes in having a short flagellum [4], and, therefore, reduced mobility. Some motility appears to be essential for midgut binding, as dead non-motile promastigotes were not able to attach. The lack of binding also raises the possibility of the sand fly binding receptors being restricted to the base of the microvilli where they cannot be reached by the short procyclic flagella and/or a mechanical interaction between the flagellum and microvilli being important in initiating binding.

Species-species binding comparisons

In the next series of experiments we assayed the binding of *L. infantum* to *Lu. longipalpis*, a permissive vector and natural host of this parasite in South America, in the presence of various other species of *Leishmania* (Fig. 3). In these experiments the results are displayed such that if both species of parasite bind equally well they will give a value of 50% on the y-axis. If *L. infantum* binds in

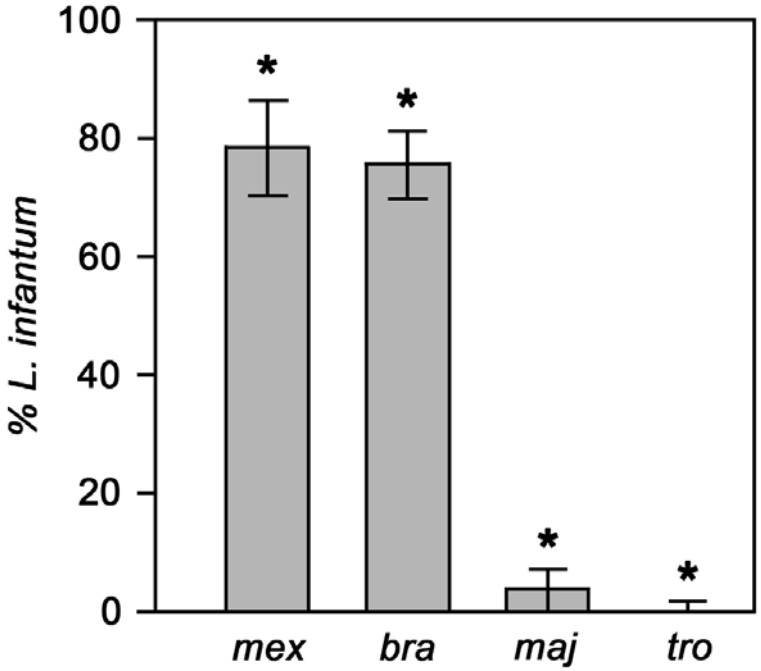


Figure 3. Binding of *L. infantum* to midguts from *Lu. longipalpis* in the presence of various *Leishmania* species (4 experiments). Competing species: *mex* = *L. mexicana*; *bra* = *L. braziliensis*; maj = *L. major*; *tro* = *L.tropica*. Bars represent the mean percentage contribution of *L. infantum* in the bound population ± S.D. and the asterisks indicate a significant difference to the *L. infantum* population (>20 dissected guts examined per experiment).
doi:10.1371/journal.pntd.0000816.g003

greater relative numbers values of greater than 50% will be found; conversely poorer binding will yield values lower than 50%. The data shows that *L. infantum* is able to bind more effectively to *Lu. longipalpis* than either *L. mexicana* or *L. braziliensis* ($P<0{\cdot}01$ in both cases), approximately 3–4 fold higher (Fig. 3). These two parasites are also found in the New World. On the other hand, *L. major* and *L. tropica* both strongly out-competed *L. infantum* in binding to *Lu. longipalpis* (Fig. 3). Binding of *L. major* accounted for over 90% of the parasites counted: not only were bound *L. major* proportionally more abundant than *L. infantum* ($P<0{\cdot}01$), but the number of *L. infantum* bound per gut was greatly reduced (from ~90 to ~20 per field), indicating that *L. major* was genuinely out-competing *L. infantum* for binding sites. This effect was even more pronounced with *L. tropica*, which completely out-competed *L. infantum* binding to *Lu. longipalpis* midguts ($P<0{\cdot}01$; ~7 per field). Such competition could never occur in nature, as *L. major* and *L. tropica* are restricted to the Old World. However, these results show that the natural parasite is not necessarily the species that can always bind *in vitro* most efficiently to the midgut of its vector.

In the final series of experiments we investigated the binding of *L. major* to midguts from various vectors, including to those of its natural host *P. papatasi* (Fig. 4A). Some of the results were as predicted from known vector-parasite combinations, but again with some interesting exceptions. *L. major* outcompeted *L. infantum* on *P. papatasi* ($P<0{\cdot}01$), this result being more expected than that found with the permissive vector reported above. Similarly, *L. major* out-competed *L. mexicana* ($P<0{\cdot}01$), which cannot complete its development in *P. papatasi*. On the other hand *L. braziliensis* was able to bind to *P. papatasi* equally well as *L. major* ($P=0{\cdot}34$), a somewhat unexpected finding given that this vector cannot support the development of *L. braziliensis* [34], or the closely related peripylarian parasite *L. panamensis* [35]. Another intriguing result was provided by *L. tropica*, which, representing about 30% of the population in competition with *L. major*, was clearly capable of *in-vitro* binding to this specific vector to a significant degree (not as well as *L. major*, though, $P<0{\cdot}01$). The competition between *L. major* and *L. tropica* was explored further in two additional vectors, *Lu. longipalpis* and *P. sergenti*. The results with *Lu. longipalpis* (Fig. 4B) were similar to those with *P. papatasi*, *L. major* being somewhat more efficient than *L. tropica* ($P<0{\cdot}01$). However, the results with *P. sergenti* were again contrary to expectations (Fig. 4C). This is the specific vector of *L. tropica* and, therefore, this might be expected to out-compete all other parasites. However, there was equal *in vitro* binding of *L. major* and *L. tropica* in *P. sergenti* ($P=0{\cdot}96$). The absolute numbers of bound promastigotes varied in these species comparisons. The lowest and highest numbers observed were for *L. infantum* and *L. major* when competing with each other in *P. papatasi*, 2–10 parasites per field to over 1000, respectively.

The ability of different sand flies to support the growth of *Leishmania* differs from species to species. Some, such as *P. papatasi* and *P. sergenti* are highly specific, and these only allow the full development of *L. major* and *L. tropica* respectively, while others are more permissive (e.g. *Lu. longipalpis* and *P. perniciosus*) and allow the development of a number of species [19]. At each stage of its development the parasite has to overcome or evade the defences of the sand fly that may prevent progression of the life-cycle. One key barrier is the ability of *Leishmania* to attach to the midgut after escaping from the peritrophic matrix, and thus greatly enhance their chances of remaining in the gut after defecation. For *P. papatasi*, the results presented here indicate a complete inability of *L. infantum* and *L. mexicana* to attach to the midgut, suggesting midgut binding is highly likely to be a crucial step in parasite establishment in this species of sand fly. Due to the competitive nature of the binding assay used, it is, however, impossible to ignore the possibility that *L. infantum* and *L. mexicana* are able to bind weakly to the midguts but are being out-competed by the *L. major* promastigotes for binding sites. Both of these interpretations are compatible with the published literature. For example, Pimenta *et al.* [7] reported binding by *L. infantum* to *P. papatasi* midguts was approximately 10% of that observed with *L. major*.

Neither *L. braziliensis* [34] nor *L. tropica* [7] are capable of completing their life-cycles within *P. papatasi*, but both bound well to *P. papatasi* guts *in vitro*, with *L. braziliensis* showing no difference in its binding abilities to those of *L. major*. A similar observation

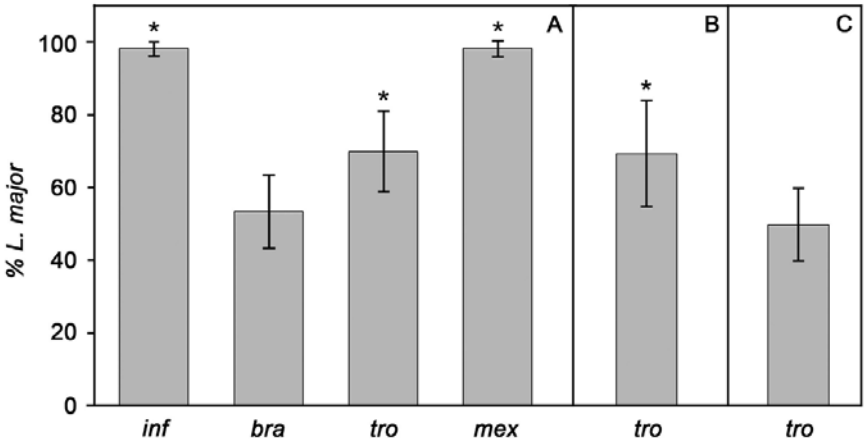


Figure 4. Binding of *L. major* to midguts from sand fly species in the presence of various *Leishmania* species (6 experiments). (A) Binding to *P. papatasi* midguts. (B) Binding to *Lu. longipalpis* midguts. (C) Binding to *P. sergenti* midguts. Competing species: inf = *L. infantum* FVI; *bra* = *L. braziliensis*; *tro* = *L. tropica*; *mex* = *L. mexicana*. Bars represent the mean percentage contribution of *L. major* in the bound population ± S.D. and the asterisks indicate a significant difference to the *L. major* population (>20 dissected guts examined per experiment).
doi:10.1371/journal.pntd.0000816.g004

was made by Warburg *et al.* [10] in studies on the flagellar attachment of *L. panamensis*. Obviously, in these cases the ability to bind to the midgut *in vitro* is not the selective refractory barrier to infection of these species in *P. papatasi*. Experimental infections with *L. panamensis* in *P. papatasi* [35] demonstrated the majority of flies lost their infection during defecation of the blood meal. That most flies retained an intact peritrophic matrix prior to defecation suggests the parasites were probably unable to break free of the blood meal and enter the midgut lumen, rendering their ability to bind to midguts irrelevant. In the few flies that did remain infected, other refractory barriers appeared to exist, such as the parasites lacking the ability to migrate to the anterior midgut, and differentiation to metacyclic promastigotes was not observed.

In contrast to *P.* papatasi, *Lu. longipalpis* supports the development of every species of *Leishmania* tested so far. A natural vector of *L. infantum* in South America, it was surprising to observe an almost complete lack of binding of this parasite when in mixed populations with *L. major* or *L. tropica*. Of course, neither of these Old World parasites are encountered by *Lu. longipalpis* in its natural range, but both bound very strongly to the fly's midgut *in vitro* and effectively outcompeted *L. infantum* for binding sites. The nature of the competition effect is unknown, but two possible scenarios are: the parasites are binding to the same receptors, with *L. major* simply having a higher affinity for the receptor; or each species is binding to distinct receptors, but the binding of *L. major* is much stronger and steric hindrance is preventing *L. infantum* from accessing their receptors.

The results with *P. sergenti* also contradict those reported by Kamhawi *et al.* [20] who observed a much lower binding of *L. major* to *P. sergenti* guts in comparison to *L. tropica*. It is possible these apparently conflicting results may have arisen from differences in the experimental design of the binding assays. Two major differences that may account for the higher binding observed here are incubation in different media at different pH and a longer incubation time. Because gut tissue is unstable in PBS after only 10 minutes incubation [20], we decided to use GIM since it is more physiologically similar to the conditions found inside the fly gut and it causes no noticeable changes in the tissue stability for at least 4 hours. The increased health of the gut tissue may alone be enough to account for the differences between the results, but the increase in incubation times from 10 minutes to the 45 minutes used here may also have contributed to a higher observed attachment.

In conclusion, the results of this study confirm that the ability of the *Leishmania* parasite to bind to the midgut epithelium is important and probably essential to transmission. In assessing this interaction it is very important to use the correct life-cycle stages, otherwise erroneous conclusions may be drawn. *In vitro* cultures usually contain a mixture of stages, but can easily lack nectomonad and/or leptomonad promastigotes, depending on culture methodology and time of usage with respect to growth phase (early/mid/late exponential or stationary phase). Such populations will, therefore, lack some or all of the capacity to bind to sand fly midguts. However, whilst binding is necessary it is clearly not sufficient to explain vector specificity and other factors must play a role in limiting the development of parasites in certain vectors. For example, the binding of inappropriate species may not be sufficiently strong in vivo to withstand midgut peristalsis. Conversely, additional factors beyond physical retention may permit the persistence of appropriate parasites in their vectors, and there is the likelihood that strain variation in binding affinity will contribute to observed vector-parasite competence in the field. Other parasite factors important for transmission are the ability of parasites to damage the stomodeal valve [36,37] and create a blocked fly through secretion of promastigote secretory gel (PSG), a prerequisite for transmission by regurgitation [4,32]. The fine tuning of parasite development to the digestive physiology of the sand fly is also likely to be one factor, as failure to escape from the peritrophic matrix in time would prevent establishment whether or not a parasite is capable of binding to the midgut. Sand fly factors include the timing and mixture of digestive enzymes produced, trypsin being one known example that exerts an antiparasitic effect [36,38,39], as well as the almost completely unknown role of the sand fly immune response in limiting or preventing parasite development [40].

Beyond these it is clear that additional factors affect the *Leishmania* species transmitted by particular sand flies in nature, including the feeding habits of the flies and prevalence of reservoir hosts. From the molecular to the ecological these are all important issues to understand individually and collectively when confronted with the changing epidemiology of leishmaniasis and new or re-emerging foci of disease.

Author Contributions

Conceived and designed the experiments: RW MDB PAB. Performed the experiments: RW MDB AD LJ. Analyzed the data: RW MDB AD PV PAB. Contributed reagents/materials/analysis tools: AD LJ RJD PV. Wrote the paper: RW PV PAB.

References

1. Lawyer PG, Young DG, Butler JF, Akin DE (1987) Development of *Leishmania mexicana* in *Lutzomyia diabolica* and *Lutzomyia shannoni* (Diptera: Psychodidae). J Med Entomol 24: 347–355.
2. Walters LL, Modi GB, Chaplin GL, Tesh RB (1989) Ultrastructural development of *Leishmania chagasi* in its vector, *Lutzomyia longipalpis* (Diptera: Psychodidae). Am J Trop Med Hyg 41: 295–317.
3. Cihakova J, Volf P (1997) Development of different *Leishmania major* strains in vector sandflies *Phlebotomus papatasi* and *P. duboscqi*. Ann Trop Med Parasitol 91: 267–279.
4. Rogers ME, Chance ML, Bates PA (2002) The role of promastigote secretory gel in the origin and transmission of the infective stage of *Leishmania mexicana* by the sand fly *Lutzomyia longipalpis*. Parasitol 124: 495–507.
5. Gossage SM, Rogers ME, Bates PA (2003) Two separate growth phases during the development of *Leishmania* in sand flies: implications for understanding the life cycle. Int J Parasitol 33: 1027–1034.
6. Pimenta PFP, Turco SJ, McConville MJ, Lawyer PG, Perkins PV, et al. (1992) Stage-specific adhesion of *Leishmania* promastigotes to the sandfly midgut. Science 256: 1812–1815.
7. Pimenta PFP, Saraiva EMB, Rowton E, Modi GB, Garraway LA, et al. (1994) Evidence that the vectorial competence of phlebotomine sand flies for different species of *Leishmania* is controlled by structural polymorphisms in the surface lipophosphoglycan. Proc Natl Acad Sci USA 91: 9155–9159.
8. Sacks DL, Saraiva EM, Rowton E, Turco SJ, Pimenta PF (1994) The role of the lipophosphoglycan of *Leishmania* in vector competence. Parasitol 108: S55–S62.
9. Sacks DL, Modi G, Rowton E, Spath G, Epstein L, et al. (2000) The role of phosphoglycans in *Leishmania*-sand fly interactions. Proc Natl Acad Sci USA 97: 406–411.
10. Warburg A, Tesh RB, McMahon-Pratt D (1989) Studies on the attachment of *Leishmania* flagella to sand fly midgut epithelium. J Protozool 36: 613–617.
11. Lang T, Warburg A, Sacks DL, Croft SL, Lane RP, et al. (1991) Transmission and scanning EM-immunogold labelling of *Leishmania major* lipophosphoglycan in the sand fly *Phlebotomus papatasi*. Eur J Cell Biol 55: 362–372.
12. Butcher BA, Turco SJ, Hilty BA, Pimenta PF, Panunzio M, et al. (1996) Deficiency in β1,3-galactosyltransferase of a *Leishmania major* lipophosphoglycan mutant adversely influences the *Leishmania*-sand fly interaction. Proc Natl Acad Sci USA 271: 20573–20579.
13. Kamhawi S, Ramalho-Ortigao M, Pham VM, Kumar S, Lawyer PG, et al. (2004) A role for insect galectins in parasite survival. Cell 119: 329–341.
14. Sacks D, Kamhawi S (2001) Molecular aspects of parasite-vector and vector-host interactions in Leishmaniasis. Ann Rev Microbiol 55: 453–483.
15. Myskova J, Svobodova M, Beverley SM, Volf P (2007) A lipophophosglycan-independent development of *Leishmania* in permissive sand flies. Microb Infect 9: 317–324.
16. Svarovska A, Ant TH, Seblova V, Jecna L, Beverley SM, et al. (2010) *Leishmania major* glycosylation mutants require phosphoglycans (lpg2-) but not lipophosphoglycan (lpg1-) for survival in permissive sand fly vectors. PLoS Neg Trop Dis 4: e580.
17. Sacks DL (2001) *Leishmania*-sand fly interactions controlling species-specific vector competence. Cellul Microbiol 3: 189–196.

18. Kamhawi S (2006) Phlebotomine sand flies and *Leishmania* parasites: friends or foes? Trends Parasitol 22: 439–445.
19. Volf P, Myskova J (2007) Sand flies and *Leishmania*: specific versus permissive vectors. Trends Parasitol 23: 91–92.
20. Kamhawi S, Modi GB, Pimenta PFP, Rowton E, Sacks DL (2000) The vectorial competence of *Phlebotomus sergenti* is specific for *Leishmania tropica* and is controlled by species-specific, lipophosphoglycan-mediated midgut attachment. Parasitol 121: 25–33.
21. Svobodova M, Votypka J, Peckova J, Dvorak V, Nasseredin A, et al. (2006) Distinct transmission cycles of *Leishmania tropica* in two adjacent foci, Nothern Israel. Emerg Infect Dis 12: 1860–1868.
22. Walters LL, Irons KP, Chaplin G, Tesh RB (1993) Life cycle of *Leishmania major* (Kinetoplastida: Trypanosomatidae) in the neotropical sand fly *Lutzomyia longipalpis* (Diptera: Psychodidae). J Med Entomol 30: 699–718.
23. Volf P, Benkova I, Myskova J, Sadlova J, Campino L, et al. (2007) Increased transmission potential of *Leishmania major/Leishmania infantum* hybrids. Int J Parasitol 37: 589–593.
24. Sadlova J, Hajmova M, Volf P (2003) *Phlebotomus* (*Adlerius*) *halepensis* vector competence for *Leishmania major* and *L. tropica*. Med Vet Entomol 17: 244–250.
25. Walters LL, Modi GB, Tesh RB, Burrage T (1987) Host-parasite relationship of *Leishmania mexicana mexicana* and *Lutzomyia abonnenci* (Diptera: Psychodidae). Am J Trop Med Hyg 36: 2 94–314.
26. Sacks DL, Pimenta PFP, McConville MJ, Schneider P, Turco SJ (1995) Stage-specific binding of *Leishmania donovani* to the sand fly vector midgut is regulated by conformational changes in the abundant surface lipophosphoglycan. J Exp Med 181: 685–697.
27. Soares RPP, Cardosa TL, Barron T, Araujo MSS, Pimenta PFP, et al. (2005) *Leishmania braziliensis*: a novel mechanism in the lipophosphoglycan regulation during metacyclogenesis. Int J Parasitol 35: 245–253.
28. Hendricks LD, Wood DE, Hadjuk ME (1978) Haemoflagellates: commercially available liquid media for rapid cultivation. Parasitol 76: 309–316.
29. Soares RPP, Macedo ME, Ropert C, Gontijo NF, Almeida IC, et al. (2002) *Leishmania chagasi*: lipophosphoglycan characterization and binding to the midgut of the sand fly vector *Lutzomyia longipalpis*. Mol Biochem Parasitol 121: 213–224.
30. Sacks DL, Brodin TN, Turco SJ (1990) Developmental modification of the lipophosphoglycan from *Leishmania major* promastigotes during metacyclogenesis. Mol Biochem Parasitol 42: 225–234.
31. McConville MJ, Turco SJ, Ferguson MAJ, Sacks DL (1992) Developmental modification of lipophosphoglycan during the differentiation of *Leishmania major* promastigotes to an infectious stage. EMBO J 11: 3593–3600.
32. Rogers ME, Ilg T, Nikolaev AV, Ferguson MAJ, Bates PA (2004) Transmission of cutaneous leishmaniasis by sand flies is enhanced by regurgitation of fPPG. Nature 430: 463–467.
33. Saraiva EMB, Pimenta PFP, Brodin TN, Rowton E, Modi GB, et al. (1995) Changes in lipophosphoglycan and gene expression associated with the development of *Leishmania major* in *Phlebotomus papatasi*. Parasitol 111: 275–287.
34. Adler S, Theodor O (1927) The behaviour of cultures of *Leishmania tropica*, *L. infantum*, and *L. braziliense* in the sandfly, *Phlebotomus papatasi*. Nature 119: 48–49.
35. Walters LL, Irons KP, Modi GB, Tesh RB (1992) Refractory barriers in the sand fly *Phlebotomus papatasi* (Diptera: Psychodidae) to infection with *Leishmania panamensis*. Am J Trop Med Hyg 46: 211–228.
36. Rogers ME, Hajmová M, Joshi MB, Sadlova J, Dwyer DM, et al. (2008) *Leishmania* chitinase facilitates colonization of sand fly vectors and enhances transmission to mice. Cellul Microbiol 10: 1363–1372.
37. Volf P, Hajmova M, Sadlova J, Votypka J (2004) Blocked stomodeal valve of the insect vector: similar mechanism of transmission in two trypanosomatid models. Int J Parasitol 34: 1221–1227.
38. Pimenta PF, Modi GB, Pereira ST, Shahabuddin M, Sacks DL (1997) A novel role for the peritrophic matrix in protecting *Leishmania* from the hydrolytic activities of the sand fly midgut. Parasitol 115: 359–369.
39. Sant'Anna MRV, Diaz-Albiter H, Mubaraki M, Dillon RJ, Bates PA (2009) Inhibition of trypsin expression in *Lutzomyia longipalpis* using RNAi enhances the survival of *Leishmania*. Parasites Vectors 2: 62.
40. Boulanger N, Lowenberger C, Volf P, Ursic R, Sigutova L, et al. (2004) Characterization of a defensin from the sand fly *Phlebotomus duboscqi* induced by challenge with bacteria or the protozoan parasite *Leishmania major*. Infect Immun 72: 7140–7146.

OPEN ACCESS Freely available online

PLoS NEGLECTED TROPICAL DISEASES

Leishmania major Glycosylation Mutants Require Phosphoglycans (*lpg2*$^{-}$) but Not Lipophosphoglycan (*lpg1*$^{-}$) for Survival in Permissive Sand Fly Vectors

Anna Svárovská[1], Thomas H. Ant[2], Veronika Seblová[1], Lucie Jecná[1], Stephen M. Beverley[3], Petr Volf[1]*

1 Department of Parasitology, Faculty of Science, Charles University in Prague, Czech Republic, **2** Keele University, Keele, Staffordshire, United Kingdom, **3** Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri, United States of America

Abstract

Background: Sand fly species able to support the survival of the protozoan parasite *Leishmania* have been classified as permissive or specific, based upon their ability to support a wide or limited range of strains and/or species. Studies of a limited number of fly/parasite species combinations have implicated parasite surface molecules in this process and here we provide further evidence in support of this proposal. We investigated the role of lipophosphoglycan (LPG) and other phosphoglycans (PGs) in sand fly survival, using *Leishmania major* mutants deficient in LPG (*lpg1*$^{-}$), and the phosphoglycan (PG)-deficient mutant *lpg2*$^{-}$. The sand fly species used were the permissive species *Phlebotomus perniciosus* and *P. argentipes*, and the specific vector *P. duboscqi*, a species resistant to *L. infantum* development.

Principal Findings: The *lpg2*$^{-}$ mutants did not survive well in any of the three sand fly species, suggesting that phosphoglycans and/or other *LPG2*-dependent molecules are required for parasite development. *In vitro*, all three *L. major* lines were equally resistant to proteolytic activity of bovine trypsin, suggesting that sand fly-specific hydrolytic proteases or other factors are the reason for the early *lpg2*$^{-}$ parasite killing. The *lpg1*$^{-}$ mutants developed late-stage infections in two permissive species, *P. perniciosus* and *P. argentipes*, where their infection rates and intensities of infections were comparable to the wild type (WT) parasites. In contrast, in *P. duboscqi* the *lpg1*$^{-}$ mutants developed significantly worse than the WT parasites.

Conclusions: In combination with previous studies, the data establish clearly that LPG is not required for *Leishmania* survival in permissive species *P. perniciosus* and *P. argentipes* but plays an important role in the specific vector *P. duboscqi*. With regard to PGs other than LPG, the data prove the importance of *LPG2*-related molecules for survival of *L. major* in the three sand fly species tested.

Citation: Svárovská A, Ant TH, Seblová V, Jecná L, Beverley SM, et al. (2010) *Leishmania major* Glycosylation Mutants Require Phosphoglycans (*lpg2*$^{-}$) but Not Lipophosphoglycan (*lpg1*$^{-}$) for Survival in Permissive Sand Fly Vectors. PLoS Negl Trop Dis 4(1): e580. doi:10.1371/journal.pntd.0000580

Editor: Yara M. Traub-Cseko, Instituto Oswaldo Cruz, Fiocruz, Brazil

Received June 22, 2009; **Accepted** November 30, 2009; **Published** January 12, 2010

Funding: The study was supported by the Ministry of Education of the Czech Republic (MSM 0021620828, LC 06009), Czech Science Foundation (GACR 206/09/0777), Wellcome Trust, and US National Institutes of Health grant AI31078 (SB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: volf@cesnet.cz

Introduction

The distribution of diseases caused by the protozoan parasite *Leishmania* is limited by the distribution of the sand fly vectors and their capacity to support parasite development. Survival of *Leishmania* parasites during bloodmeal digestion and their attachment to the midgut epithelium have been identified as two critical steps determining the vector competence. Based upon experimental tests of their ability to support development of wide or limited range of *Leishmania* species, sand flies have been classified as permissive or specific vectors [1]. According to previous investigations, there is a close evolutionary fit between *Phlebotomus papatasi* and *P. sergenti* with *Leishmania major* and *L. tropica* respectively, as other *Leishmania* species survive poorly in these sand fly hosts. In contrast, other sand flies tested (*P. argentipes*, *P. halepensis*, *P. arabicus* and *Lutzomyia longipalpis*) were broadly permissive to the development of different *Leishmania* parasites. This classification is based on experimental studies and does not imply the constraints of natural transmissions (vector capacity). However, it reflects the vector competence of permissive sand flies for transmission of various parasites (for review see [1,2]).

Leishmania surface molecules have been strongly implicated in parasites survival within sand fly vectors. *Leishmania* promastigotes synthesise an abundance of glycoconjugates composed of polymeric units based upon a conserved Gal-Man-P phosphoglycan (PG) repeating unit for review see [3]. These include the membrane-attached glycosylphosphatidyl-inositol (GPI) anchored lipospophoglycan (LPG) and proteophosphoglycan (PPG), as well as secreted forms of PPGs and secretory acid phosphatases (sAPs). PGs have been implicated in the early survival of *L. donovani* within the bloodfed midgut [4], presumably by conferring resistance to, or by modulating the activity of digestive enzymes. The role of these molecules in sand fly interactions has been studied by biochemical methods using purified LPG, PPG and other

Author Summary

Phlebotomine sand flies are small blood-feeding insects, medically important as vectors of protozoan parasites of the genus *Leishmania*. Sand flies species can be divided roughly into two groups, termed specific or permissive, depending on their ability to support development of one or a few strains vs. a broad spectrum of these parasites. In this study, we explored the ability of two *Leishmania major* glycocalyx mutants to survive within these different types of vectors. The *lpg1*$^-$ mutant, which specifically lacks lipophosphoglycan (LPG), was able to survive normally in two permissive species, *Phlebotomus argentipes* and *P. perniciosus*, but was only able to survive within the specific species *P. duboscqi* for a limited time prior to dissolution of the peritrophic matrix. Consistent with its classification as a specific sand fly vector, *P. duboscqi* was not able to support development of *L. infantum*. The *lpg2*$^-$ *L. major* mutant, which is a broader mutant and lacks all phosphoglycans including LPG and proteophosphoglycans, was unable to survive in all the three vector species tested. This study extends the knowledge on the role of *Leishmania major* surface glycoconjugates to development in three important vector species and gives supporting evidence for the existence of an LPG-independent mechanism for survival in sand flies, as well as the importance of *LPG2*-dependent glycoconjugates in parasite survival.

molecules, although the shared PG determinants across molecules makes the assignment of function to specific PG classes problematic [5,6]

Thus, complementary to studies using purified PGs, researchers have generated and studied the behaviour of mutants lacking specific LPG or PG biosynthetic genes, which in turn affect the synthesis of subsets of PG-bearing molecules. The *lpg1*$^-$ mutant lacks the *LPG1* gene (GenBank accession no. AF234766) which encodes a galactofuranosyltransferase required for synthesis of the LPG glycan core, rendering such mutants specifically deficient in LPG alone [7]. In contrast, the *Leishmania major lpg2*$^-$ mutant lacks the *LPG2* gene (GenBank accession no. AF350492) encoding a Golgi GDP-sugar transporter. As a result, these parasites fail to synthesize LPG and all other PGs [8,9]. Recent studies suggest that the *lpg2*$^-$ mutant may lack additional, as yet unidentified and likely rare glycoconjugates [10]. Importantly, the LPG and virulence phenotypes of the *L. major lpg1*$^-$ and *lpg2*$^-$ lines in both mammalian and sand fly infections were restored to WT following re-expression of the cognate gene, confirming the genetic specificity of the virulence defects [4,5,7,9–13].

The genetic and biochemical approaches above have established that LPG, the dominant surface glycoconjugate of *Leishmania* promastigotes, mediates attachment to the midgut epithelium in *P. papatasi*, preventing the loss *L. major* parasites during blood meal excretion, as *lpg1*$^-$ parasites survive the initial stages of fly infection but are subsequently lost due to a failure to bind to a sand fly midgut LPG receptor [4,14,15]. Notably *P. papatasi* is considered to be a highly specific vector, in that other species of *Leishmania* are unable to establish infection in this sand fly [16–19]. In contrast, in *P. arabicus* and *Lu. longipalpis*, two species shown to be permissive vectors by virtue of their ability support development of various *Leishmania* species, LPG-deficient *lpg1*$^-$ *L. major* develop and produce mature infections [20]. From these data Myskova *et. al* hypothesised that LPG is required in specific vectors, while in permissive vectors *Leishmania* bind via an LPG independent mechanism.

In this work we further test this hypothesis by infections of three additional sand fly species with WT and mutant *L. major*. In addition, we attempted to assess the importance of LPG and other *LPG2*-dependent molecules in protection against proteolytic attack by exposing the mutant parasite lines to the action of bovine trypsin *in vitro*. Importantly, the three sand fly species used in the study are important vectors known to transmit *Leishmania*. *Phlebotomus duboscqi* is a vector of cutaneous leishmaniasis caused by *L. major* in sub-Saharan Africa [21,22]. It is a sister species of *P. papatasi* and belongs to the same subgenus. Unlike *P. papatasi*, some populations of *P. duboscqi* have been shown experimentally to support development of *L. tropica* [16]. Midgut glycosylation and the degree of permissivity of this species are unclear. We addressed the question of permissivity of *P. duboscqi* sand flies in this study by infecting them with *L. infantum*. The other two species used are permissive vectors transmitting parasites of *L. donovani* complex. Myskova *et al.* [20] demonstrated that both, *P. argentipes* and *P. perniciosus* posses midgut glycoproteins with HPA (*Helix pomatia* agglutinin, lectin with specificity to N-acetyl-D-galactosamine)-binding epitopes. *Phlebotomus argentipes* is a vector of visceral anthroponotic leishmaniasis caused by *Leishmania donovani* in the Indian subcontinent [23]. In experimental conditions it supports development of *L. donovani*, *L. amazonensis*, *L. major* and *L. tropica* [4,19,24]. *Phlebotomus perniciosus* is a vector of *Leishmania infantum* in the western Mediterranean and in experimental conditions it supports the development of *L. tropica* (V.S. and P.V., unpublished results).

Materials and Methods

Parasites

Leishmania infantum MHOM/TR/2000/OG-VL and three lines of *Leishmania major* LV39 clone 5 (MRHO/SU/1959/Neal P) [25] were used in this work. The *L. major LPG1* and *LPG2* knockout mutants *lpg1*$^-$ and *lpg2*$^-$ were generated in the LV39 clone 5 background previously [7,9]. Parasites were maintained at 23°C on medium 199 supplemented with 20% foetal calf serum (Gibco) and gentamicin (50μg/ml). For the mutated lines, selection antibiotics were added to the culture medium as follows: hygromycin B (15μg/ml) for the *lpg2*$^-$ mutant; hygromycin (15 μg/ml) and puromycin (11μg/ml) for the *lpg1*$^-$ mutant. Prior to sand fly infections, parasites were washed by centrifugation and resuspended in saline solution.

Sand fly colonies

Laboratory colonies of three sand fly species were used: *Phlebotomus perniciosus* (originally from Spain), *P. argentipes* (originally from India) and *P. duboscqi* (originally from Senegal). Colonies were maintained in conditions described previously [26]. Adults were maintained at 26°C and fed on 50% sucrose *ad libitum*.

Leishmania development in sand flies

Female sand flies (5–10 days old) were fed through a chick skin membrane with 4–5 day old promastigotes at cell density of 5×10^6 (*P. duboscqi* infections with *L. major*) or 1×10^6 promastigotes/ml (all other infections, including *P. duboscqi* with *L. infantum*) in heat inactivated rabbit blood (Bioveta, Ivanovice). Blood-engorged females were maintained at 26°C with access to cotton wool soaked in 50% solution of sugar in distilled water and sacrificed for microscopical examination and counting of parasites in the midgut 2 and 5 or 9 days post infection. Intensity of infection was graded as light (<100 parasites/gut), moderate (100–1000 parasites/gut) or heavy (>1000 parasites/gut) as described previously [27]. *Phlebotomus perniciosus* and *P. duboscqi* defecate between 75 and

95 hours post-feeding [28] and *L. major* colonized their stomodeal valve on days 7–9 post-feeding [27]. Preliminary experiments showed that *Phlebotomus argentipes* defecates 2–3 days post-feeding and parasites reached the stomodeal valve by day 5 already.

Experiments were repeated twice. The χ^2 test was used for comparison of infection rates (number of infected versus uninfected females) and intensities of infection (heavy, moderate light, zero) between the WT and the mutant lines using S-PLUS 2000 programme.

Parasite susceptibility to bovine trypsin

Promastigotes of a 4-day culture were washed in medium 199 (M 199), adjusted to the concentration of 3×10^6 cells/ml of M 199 and exposed to bovine trypsin (13.500 BAEE units/ml) (Sigma) alone or bovine trypsin plus 6% human haemoglobin (Sigma; one BAEE unit will produce a $_{\Delta}A^{253}$ of 0.001 per min at pH 7.6 at 25 °C using benzoyl-L-arginine ethyl ester (BAEE) as substrate). In control groups, parasites were cultivated in M 199 alone. After 24 hours at 23°C, parasite numbers were determined by haemocytometer counting. Assay was performed in triplicate and the experiment was repeated twice. Data were evaluated statistically by means of ANOVA test.

Detection of glycoconjugates in *P. duboscqi* midgut lysates

Midguts of female *P. duboscqi* were homogenized in Tris buffer (20mM Tris, 150mM NaCl, pH 7.6) and proteins were analyzed by SDS PAGE (10% gel, reducing conditions, 10 µg protein per lane) followed by western blotting. The nitrocellulose membrane was incubated in Tris buffer with 0.05% Tween 20 (Tris-Tw) with 5% bovine serum albumin overnight and then with biotinylated lectins (Sigma) in Tris-Tw with 1% BSA in the following concentrations: concanavalin A (Con, 2.5 µg/ml), *Pisum sativum* agglutinin (PSA, 2.5 µg/ml), *Helix pomatia* agglutinin (HPA, 1 µg/ml), *Ricinus communis* agglutinin (RCA, 0.5 µg/ml), Soybean agglutinin (SBA, 10 µg/ml). After repeated washing the blots were incubated with streptavidin peroxidase (2.5 µg/ml) in Tris-Tw and developed in 3,3′-diaminobenzidine solution. The specificity of Con A and PSA reactions were controlled by addition of 250mM methyl-mannopyranoside as an inhibitory sugar.

Results

Development of *L. major* in *P. duboscqi*

Phlebotomus duboscqi sand flies were infected with WT, *lpg1*$^-$ and *lpg2*$^-$ mutants of *L. major* in order to study the role of LPG and other PGs. On day 2 post-infection, no differences were observed between development of WT and the *lpg1*$^-$ line, with both showing very high rates of infection (97% and 93% respectively), with about 75% of heavy infections (Fig. 1). The *lpg2*$^-$ mutant survived less well however; their infection rate was significantly lower (74%; $P<0.01$), with only 24% of heavy infections.

In contrast to day 2, on day 9 post infection the *lpg1*$^-$ mutant survived much more poorly than WT, with only 25% of the flies retaining parasites, as opposed to 92% for WT ($P<0.01$). Moreover, the *lpg1*$^-$ line produced very few mature infections colonizing the stomodeal valve in only 10% of females. Notably the *lpg2*$^-$ line did not survive, as no *lpg2*$^-$ parasites were found in the midguts by day 9 (Fig. 1). Previous studies have established that phenotypes arising from the *lpg1*$^-$ and *lpg2*$^-$ mutants in diverse settings are specific, as they are reversed by complementation with the cognate gene, and thus do not arise as a result of nonspecific culture passage or other sources [4,7,9,13]. Thus these mutant

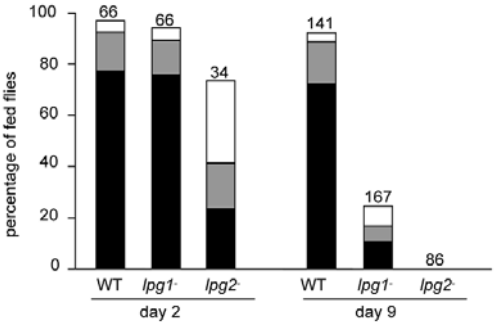


Figure 1. Development of *L. major* mutants in *P. duboscqi*. *Phlebotomus duboscqi* females were infected with *Leishmania major* wild type (WT) or mutants lacking LPG (*lpg1*$^-$) or all *LPG2*-dependent molecules (*lpg2*$^-$). Day 2 - dissection before defecation (48hours post-infection), day 9 - dissection after defecation. Infections were classified into three cathegories: heavy (more than 1000 promastigotes per gut) - black bars, moderate (100–1000) - grey bars, light (1–100) - white bars. Numbers above the bars indicate the number of dissected females.
doi:10.1371/journal.pntd.0000580.g001

data argue that, as seen previously in the specific sand fly *P. papatasi*, LPG is required for late but not early survival [4], while *LPG2* is important for early survival and essential for late stage survival.

Development of *L. major* lines in *P. argentipes*

Similar to the results obtained in *P. duboscqi*, on day 2 no statistically significant differences were found between the WT and the *lpg1*$^-$ mutant, while the *lpg2*$^-$ mutant was severely impaired (**Fig. 2**). Very high infection rates (96%) were present in both the *WT* and *lpg1*$^-$ lines, with heavy infections in 70% and 50%, respectively. In contrast, *lpg2*$^-$ mutants were more severely affected, with infections seen in 62% of flies but with very low parasite loads (less than 100 *Leishmania*, except for 1 fly). The

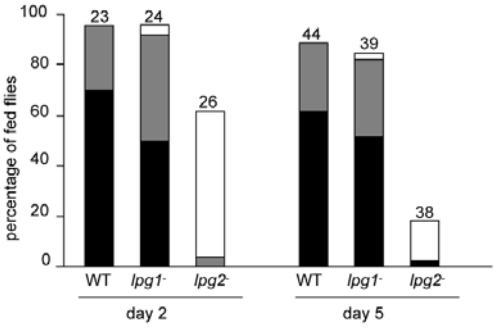


Figure 2. Development of *L. major* mutants in *P. argentipes*. *Leishmania major* lines tested were the same as in Fig. 1. Day 2 - dissection before defecation (48 hours post-infection); day 5 - dissection after defecation. The intensity of infections was evaluated as described in Fig. 1. Numbers above the bars indicate the number of dissected *P. argentipes* females.
doi:10.1371/journal.pntd.0000580.g002

differences in infection intensity and rate between the *lpg2*$^-$ and the WT were statistically highly significant ($P<0.01$).

Day-5 dissections revealed a continuation of the trends described above during late stage infections. Females infected with WT and *lpg1*$^-$ mutants retained high infection rate and high parasite loads (no statistically significant differences between the lines). Also the localizations of parasites were remarkably similar; WT and *lpg1*$^-$ parasites reached the stomodeal valve in 46% and 48%, respectively. In contrast, *lpg2*$^-$ line showed a remarkable decrease in infection rate, with only 18% of positive females (highly significant difference from the WT parasites, $P<0.01$) and no colonization of the stomodeal valve.

Development of *L. major* lines in *P. perniciosus*

On day 2, no significant differences were observed between the three lines, all of them survived well inside the peritrophic sac producing heavy infections in about 25% of females.

On day 9, *lpg2*$^-$ mutants were eliminated while *lpg1*$^-$ mutants developed similarly to the WT parasites (**Fig. 3**). WT and *lpg1*$^-$ lines developed mature infections colonizing the stomodeal valve with high parasite burdens in majority of females. In contrast, none of the *lpg2*$^-$ parasites were able to persist until day 9, suggesting that they were lost during defecation.

P. duboscqi is refractory to *L. infantum*

In order to test the degree of permissivity of *P. duboscqi*, females of this species were infected with *L. infantum*, a parasite that is not transmitted by *P. duboscqi* in nature. Promastigotes were able to survive inside the peritrophic sac during digestion of the blood-meal but they were not able to persist beyond defecation of the blood remnants. On day 2 post infective bloodmeal, 91% of the flies (11 of 12) were *Leishmania* positive while on day 8, no parasites were found in any female tested (n = 15) (data not shown). These results demonstrate that *P. duboscqi* is refractory to *L. infantum*.

P. duboscqi midgut glycosylation

As detected by western blotting with lectins, *P. duboscqi* midgut lysate displays molecules that bind Con A and PSA, lectins detecting terminal mannose residues of glycans (**Fig. 4**). Controls with inhibitory sugar (250mM methyl-mannopyranoside) confirmed the specificity of lectin reactions (data not shown). In contrast, HPA, RCA and SBA reactions were negative indicating

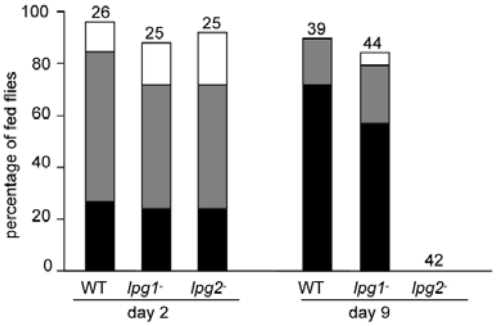


Figure 3. Development of *L. major* mutants in *P. perniciosus*. *Leishmania major* lines tested and evaluation of infections were the same as described in Fig. 1. Numbers above the bars indicate the number of dissected *P. perniciosus* females.
doi:10.1371/journal.pntd.0000580.g003

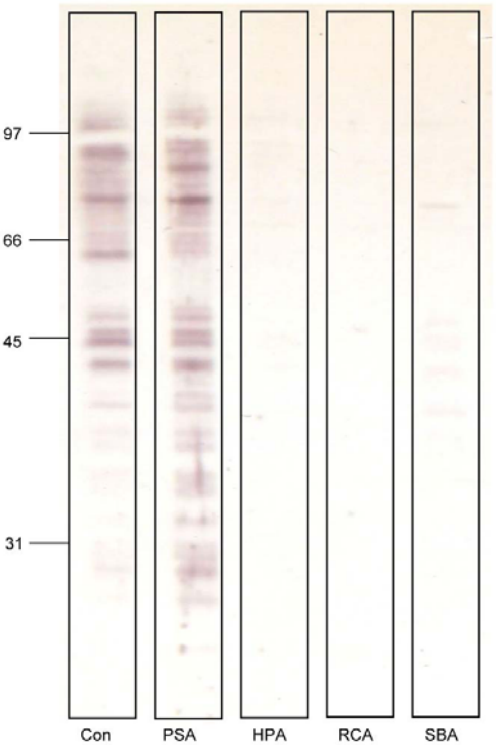


Figure 4. Western blot of *P. duboscqi* midgut proteins incubated with various lectins. Lectins concanavalin A (Con) and *Pisum sativum* agglutinin (PSA) specifically reacted with *P. duboscqi* midgut lysate. Reaction of lectins *Helix pomatia* agglutinin (HPA), *Ricinus communis* agglutinin (RCA) and Soybean agglutinin (SBA) were negative.
doi:10.1371/journal.pntd.0000580.g004

absence of β-galactose or N-acetyl-D-galactosamine residues in the midgut glycoproteins (**Fig. 4**). The lectin binding profile is similar to that previously observed in specific sand fly vectors *P. papatasi* and *P. sergenti*. In contrast, midgut lysates of all permissive sand fly species tested to date contain N-acetyl-D-galactosamine displaying glycoconjugates as detected by lectin affinity blotting [20].

Effect of bovine trypsin on *Leishmania*

The *in-vitro* growth of *L. major* promastigotes of WT, *lpg1*$^-$ and *lpg2*$^-$ lines in M 199 medium was not affected either by bovine trypsin (13.500 BAEE units/ml) or by bovine trypsin plus 6% human hemoglobin.

No significant differences in numbers of viable parasites were observed between the groups in any of the three lines tested (P = 0.84).

Discussion

The classic studies of Sacks and co-workers established a paradigm for the role of LPG in the survival of *L. major* and *L. donovani* in sand flies, mediated through binding of LPG to the sand

fly midgut [4,19]. In the case of *L. major* this interaction is now known to be mediated by the *P. papatasi* lectin PpGal ([15]). However in 2007 we reported the occurrence of LPG-independent sand fly survival, importantly only in sand fly species now termed 'permissive' as defined by their ability in experimental tests to support the development of a wide range of *Leishmania* species. In contrast, previous studies of LPG-dependency were now associated with 'selective' sand fly species (again defined by experimental tests as supporting the development of a narrow range of *Leishmania* species and/or isolates) ([20]). Given the implications of this new paradigm, it was important to garner additional data testing its validity by examining additional permissive and selective sand fly species, using the same well characterized LPG mutants studied previously. Additionally we have culled the literature and summarized the available data pertinent to this model (Table 1).

We found that *L. major* mutants specifically lacking LPG remain able to develop in the permissive vectors *P. perniciosus* and *P. argentipes* at levels resembling those of wild type parasites, with full midgut development and colonization of the stomodeal valve. These data suggest that in *P. perniciosus* and *P. argentipes* the LPG is required neither for parasite protection against digestive enzymes nor for midgut binding. LPG-independent development was previously reported for *L. major* in the permissive sand fly vectors *Lu. longipalpis* and *P. arabicus* [20], and *L. mexicana* development in *Lu. longipalpis* [29]. These data confirm and extend the results obtained in sand fly infections with all LPG-deficient *Leishmania* carried out by various laboratories to date (**Table 1**). Collectively these data provide strong support for the role of LPG in specific but not permissive sand fly vectors.

Within this data set we could only identify one potential exception, involving a study of the permissive vector *P. argentipes*, where the *L. donovani* LPG-deficient mutant line known as the 'R2D2', also defective in *LPG1* expression [30], did not survive when examined on day-5 post-infection [19]. Notably R2D2 was obtained following heavy mutagenesis and selection for LPG-deficiency, unlike the *lpg1*$^{-}$ *L. major* which was generated following precise gene targeting procedures [31]. It is well established in the genetics literature that mutagenesis frequently results in off-target deleterious effects. Our previous work established that the phenotypic alterations in the *lpg1*$^{-}$ and *lpg2*$^{-}$ arose specifically from alterations in these genes, as restoration of *LPG1* and *LPG2* function returned the phenotype to WT [7,9]. In contrast, R2D2

Table 1. Requirement for lipophosphoglycan (LPG) and other phosphoglycans (PGs) during *Leishmania* development in various sand fly species.

Vector species (colony origin)	Leishmania species	Interpretation	Mutant lines	Infection outcome	Reference
P. papatasi (Israel)	*L. major* (natural)	LPG required in late phase	*lpg1*$^{-}$ (LPG-deficient)	Low percentage of infected flies on day 5	Sacks *et al.* [4]
P. papatasi (Turkey)	*L. major* (natural)	LPG required in late phase	*lpg1*$^{-}$ (LPG-deficient)	Low percentage of infected flies on day 8	Myskova *et al.* [20]
P. duboscqi (Mali)	*L. major* (natural)	LPG required in late phase	*lpg1*$^{-}$ (LPG-deficient)	Fewer than a 1000 parasites/gut on day 7	Secundino *et al.*, (submitted)
		PGs required from early phase	*lpg2*$^{-}$ (PG-deficient)	Impaired on day 3; no parasites on day 7	
			lpg5A^{-}*/5B*$^{-}$ (PG-deficient)	Impaired on day 3; low numbers of parasites on day 6	
P. duboscqi (Senegal)	*L. major* (natural)	LPG possibly required in late phase	*lpg1*$^{-}$ (LPG-deficient)	Comparable to WT on day 4, partially impaired on days 8 and 10	Boulanger *et al.* [32]
		LPG2-related molecules required from early phase	*lpg2*$^{-}$ (PG-deficient)	Impared on day 4, no parasites on days 8 and 10	
P. duboscqi (Senegal)	*L. major* (natural)	LPG required in late phase	*lpg1*$^{-}$ (LPG-deficient)	Comparable to *WT* on day 2; impaired on day 9	This study
		LPG2-related molecules required from early phase	*lpg2*$^{-}$ (PG-deficient)	Impaired on day 2; no parasites on day 9	
Lu. longipalpis (Brazil)	*L. major* (unnatural)	LPG-independent	*lpg1*$^{-}$ (LPG-deficient)	High level of infection on day 7	Myskova *et al.* [20]
			lpg1^{-} (LPG-deficient)	High level of infection on day 7	Secundino *et al.*, (submitted)
		LPG2-related molecules required from early phase	*lpg2*$^{-}$ (PG-deficient)	Impaired on day 3; no parasites on day 7	Secundino *et al.*, (submitted)
	L. mexicana (unnatural)	LPG-independent	*lpg1*$^{-}$ (LPG-deficient)	Mature infections on day 7–9	Rogers *et al.* [29]
P. arabicus (Israel)	*L. major* (unnatural)	LPG-independent	*lpg1*$^{-}$ (LPG-deficient)	High level of infection on day 7	Myskova *et al.* [20]
P. perniciosus (Spain)	*L. major* (unnatural)	LPG-independent	*lpg1*$^{-}$ (LPG-deficient)	High level of infection on day 9	This study
		LPG2-related molecules required from early phase	*lpg2*$^{-}$ (PG-deficient)	Comparable to *WT* on day 2; no parasites on day 9	
P. argentipes (India)	*L. major* (unnatural)	LPG-independent	*lpg1*$^{-}$ (LPG deficient)	High level of infection on day 5	Thist study
		LPG2-related molecules required from early phase	*lpg2*$^{-}$ (PG-deficient)	Impaired on day 2; low numbers of parasites on day 5	
	L. donovani (natural)	LPG possibly required	R2D2 (LPG-deficient)	Severely impaired on day 5	Pimenta *et al.* [19]; Sacks *et al.* [4]
		PGs required from early phase	C3PO (PG-deficient)	Impaired on day 2; no parasites on day 5	Sacks *et al.* [4]
			lpg2^{-} (PG-deficient)	Impaired on day 2; no parasites on day 5	Sacks *et al.* [4]

doi:10.1371/journal.pntd.0000580.t001

failed this test, as restoration of *LPG1* expression to R2D2 only weakly restored both LPG and survival in *P. argentipes* [4].

In contrast to the permissive vectors, the development of *L. major lpg1*$^-$ mutants was severely impaired in the specific vector *P. duboscqi*. Although the early infections were similar to those of the WT parasites, there was a substantial decrease in the *lpg1*$^-$ infections rate after defecation of the bloodmeal. In very few females the *lpg1*$^-$ mutants produced mature late stage infections. Our results extend those reported in the study by Boulanger *et al.* [32] performed with a small number of sand flies. Similar results with *L. major lpg1*$^-$ mutants in *P. duboscqi* were recently obtained by Secundino *et al.* (submitted) (**Table 1**). Our additional experiment confirmed that *P. duboscqi* is not permissive to *L. infantum* development and can therefore be classified as a specific vector. Moreover, lectin affinity blotting revealed that unlike *P. perniciosus* and *P. argentipes*, there are no N-acetyl-D-galactosamine- displaying epitopes in *P. duboscqi* midgut (Fig 4). Such glycoconjugates have been suggested as potential *Leishmania* ligands in the midgut of permissive vector species [20]. In conclusion, this study gives supporting evidence to the present distinction of sand flies into categories based on their susceptibility to various *Leishmania* species [1]. Together with the results of Myskova *et al.* [20] and Rogers *et al.* [29], our studies of *L. major lpg1*$^-$ development suggest the presence of a an LPG-independent parasite-binding mechanism within the midgut of permissive sand flies.

Unlike LPG-deficient *lpg1*$^-$ mutants, PG-deficient *lpg2*$^-$ mutants additionally were impaired in early development in sand fly and unable to survive at all stages in all sand fly species tested. In *P. argentipes* and *P. duboscqi* parasites of this line are severely impaired as early as day 2 post infection. For *P. duboscqi*, these data stand in line with those of Boulanger *et al.* [32]. A similar finding was described for PG-deficient mutants of *L. donovani* in *P. argentipes* by Sacks *et al.* [4]. Moreover, Secundino *et al.* (submitted) have recently made similar observations in *P. duboscqi* originating from Mali and in *Lu. longipalpis* (see **Table 1**).

It has been hypothesized that parasite death in the pre-defecated sand fly midgut is attributable to digestive enzymes and that the phosphoglycans other than LPG confer resistance to the proteolytic attack [33], specifically phosphoglycans dependent on the activity of *LPG2*. While *LPG2* has been suggested to affect synthesis of other glycoconjugates beyond PGs, this possibility was excluded through studies of a second PG-deficient mutant, defective due to a lack of the UDP-Gal transporters *LPG5*A and *LPG5B* by gene targeting, showing that it is also unable to survive the late stages of *P. duboscqi* infection (Secundino *et al.*, submitted). Thus, it is likely that the defects in *lpg2*$^-$ infection of midguts described here and previously arise primarily through loss of PGs other than LPG, potentially the PPGs common to all species, or sAPs which occur in *L. major* albeit to lesser extents than in other *Leishmania* species [34].

In this work we also attempted to assess the importance of *LPG2*-dependent molecules by exposing parasites to the action of a proteolytic enzyme *in vitro*. As trypsin-like proteases were described as the most abundant digestive enzymes in both *P. papatasi* and *Lu. longipalpis* midgut after bloodfeeding [35,36], bovine panceratic trypsin was chosen for these experiments. The bovine enzyme used shares all the conserved amino acid residues that influence the substrate specificity with sand fly midgut trypsin-like molecules. The *lpg2*$^-$ promastigotes lacking surface PGs did not prove to be more vulnerable to trypsin activity than the WT and *lpg1*$^-$ parasites whose resistance to trypsin has been previously reported [13]. These results could argue against a role of PGs in conferring resistance of promastigotes to the trypsin-like digestive enzymes in the sand fly gut. However, in light of the studies of Secundino *et al* (submitted), we think it more likely that bovine trypsin is not a good model for the activity and/or properties of all the proteolytic contents of the sand fly midgut. Most importantly, Secundino *et al* showed that inhibition of tryptic and other proteolytic activity in the midgut does in part rescue the survival of *lpg2*$^-$ parasites, although they do not exclude the possibility that also other factors contribute to the inability of the *lpg2*$^-$ to survive within the bloodmeal. Sand fly immunity has been shown to play a major role in the control of bacterial and parasitic infections [32] and potentially that molecules such as antimicrobial peptides secreted to the midgut lumen could contribute to the destruction of the *lpg2*$^-$ mutant parasite.

In summary, this study demonstrates that an LPG-independent mechanism of attachment of *Leishmania* is a feature common to permissive sand fly species. It also proves the importance of *LPG2*-dependent molecules in the survival of *L. major* in various sand fly vectors. Moreover, it brings an evidence that *P. duboscqi* is not able to support development of *L. infantum* and therefore can be classified as a specific vector.

Acknowledgments

We thank David Sacks and colleagues for sharing findings from their submitted work (Secundino *et al.*). We thank Ivana Benkova for help with the experiments.

Author Contributions

Conceived and designed the experiments: AS PV. Performed the experiments: AS THA VS LJ. Analyzed the data: AS THA VS LJ SMB PV. Contributed reagents/materials/analysis tools: SMB. Wrote the paper: AS THA SMB PV.

References

1. Volf P, Myskova J (2007) Sand flies and *Leishmania*: specific versus permissive vectors. Trends Parasitol 23: 91–92.
2. Kamhawi S (2006) Phlebotomine sand flies and Leishmania parasites: friends or foes? Trends Parasitol 22: 439–445.
3. Ilg T (2000) Proteophosphoglycans of Leishmania. Parasitol Today 16: 489–497.
4. Sacks DL, Modi G, Rowton E, Spath G, Epstein L, et al. (2000) The role of phosphoglycans in Leishmania-sand fly interactions. Proc Natl Acad Sci USA 97: 406–411.
5. Turco SJ, Spath GF, Beverley SM (2001) Is lipophosphoglycan a virulence factor? A surprising diversity between Leishmania species. Trends Parasitol 17: 223–226.
6. Beverley SM, Turco SJ (1998) Lipophosphoglycan (LPG) and the identification of virulence genes in the protozoan parasite Leishmania. Trends Microbiol 6: 35–40.
7. Spath GF, Epstein L, Leader B, Singer SM, Avila HA, et al. (2000) Lipophosphoglycan is a virulence factor distinct from related glycoconjugates in the protozoan parasite leishmania major. Proc Natl Acad Sci USA 97: 9258–9263.
8. Descoteaux A, Luo Y, Turco SJ, Beverley SM (1995) A specialized pathway affecting virulence glycoconjugates of Leishmania. Science 269: 1869–1872.
9. Spath GF, Lye LF, Segawa H, Sacks DL, Turco SJ, et al. (2003) Persistence without pathology in phosphoglycan-deficient Leishmania major. Science 301: 1241–1243.
10. Capul AA, Hickerson S, Barron T, Turco SJ, Beverley SM (2007) Comparisons of mutants lacking the Golgi UDP-galactose or GDP-mannose transporters establish that phosphoglycans are important for promastigote but not amastigote virulence in Leishmania major. Infect Immun 75: 4629–4637.
11. Uzonna JE, Spath GF, Beverley SM, Scott P (2004) Vaccination with phosphoglycan-deficient Leishmania major protects highly susceptible mice from virulent challenge without inducing a strong Th1 response. J Immunol 172: 3793–3797.
12. Capul AA, Barron T, Dobson DE, Turco SJ, Beverley SM (2007) Two functionally divergent UDP-Gal nucleotide sugar transporters participate in phosphoglycan synthesis in Leishmania major. J Biol Chem 282: 14006–14017.
13. Spath GF, Garraway LA, Turco SJ, Beverley SM (2003) The role(s) of lipophosphoglycan (LPG) in the establishment of Leishmania major infections in mammalian hosts. Proc Natl Acad Sci USA 100: 9536–9541.

14. Ilg T (2001) Lipophosphoglycan of the protozoan parasite Leishmania: stage- and species-specific importance for colonization of the sandfly vector, transmission and virulence to mammals. Med Microbiol Immunol 190: 13–17.
15. Kamhawi S, Ramalho-Ortigao M, Pham VM, Kumar S, Lawyer PG, et al. (2004) A role for insect galectins in parasite survival. Cell 119: 329–341.
16. Killick-Kendrick R, Killick-Kendrick M, Tang Y (1994) Anthroponotic cutaneous leishmaniasis in Kabul, Afghanistan: The low susceptibility of Phlebotomus papatasi Leishmania tropica. Trans R Soc Trop Med Hyg 88: 252–253.
17. Heyneman D (1963) 12. Comparison of experimental Leishmania donovani infection. Am J Trop Med Hyg 12: 725–740.
18. Walters LI, Irons KP, Modi GB, Tesh RB (1992) Refractory barriers in the sandfly Phlebotomus papatasi (Diptera: Psychodidae) to infection with Leishmania panamensis. Am J Trop Med Hyg 46: 211–228.
19. Pimenta PFP, Saraiva EMB, Rowton E, Modi GB, Garraway LA, et al. (1994) Evidence that the vectorial competence of phlebotomine sand flies for different species of Leishmania is controlled by structural polymorphisms in the surface lipophosphoglycan. Proc Natl Acad Sci USA 91: 9155–9159.
20. Myskova J, Svobodova M, Beverley SM, Volf P (2007) A lipophosphoglycan-independent development of Leishmania in permissive sand flies. Microbes Infect 9: 317–324.
21. Lawyer PG, Githure JI, Anjili CO, Olobo JO, Koech DK, et al. (1990) Experimental transmission of Leishmania major vervet monkeys (Cercopithecus aethiops) by bites of Phlebotomus duboscqi (Diptera: Psychodidae). Trans R Soc Trop Med Hyg 84: 229–232.
22. Beach R, Kiilu G, Hendricks L, Oster C, Leeuwenburg J (1984) Cutaneous leishmaniasis in Kenya - transmission of Leishmania major to man by the bite of a naturally infected Phlebotomus duboscqi. Trans R Soc Trop Med Hyg 78: 747–751.
23. Lane RP, Pile MM, Amerasinghe FP (1990) Anthropophagy and Aggregation Behavior of the Sandfly Phlebotomus-Argentipes in Sri-Lanka. Med Vet Entomol 4: 79–88.
24. Shortt HE, Smith ROA, Swaminath CS, Krishnan KV (1931) Transmission of kala-azar by the bite of Phlebotomus argentipes. Ind J Med Res 18: 1373–1375.
25. Marchand M, Daoud S, Titus RG, Louis J, Boon T (1987) Variants with reduced virulence derived from Leishmania major after mutagen treatment. Parasite Immunol 9: 81–92.
26. Benkova I, Volf P (2007) Effect of temperature on metabolism of *Phlebotomus papatasi* (Diptera : Psychodidae). Journal Med Entomol 44: 150–154.
27. Cihakova J, Volf P (1997) Development of different Leishmania major strains in the vector sandflies Phlebotomus papatasi and P-duboscqi. Ann Trop Med Parasitol 91: 267–279.
28. Volf P, Killick-Kendrick R (1996) Post-engorgement dynamics of haemagglutination activity in the midgut of phlebotomine sandflies. Med Vet Entomol 10: 247–250.
29. Rogers ME, Ilg T, Nikolaev AV, Ferguson MAJ, Bates PA (2004) Transmission of cutaneous leishmaniasis by sand flies is enhanced by regurgitation of fPPG. Nature 430: 463–467.
30. Ryan KA, Garraway LA, Descoteaux A, Turco SJ, Beverley SM (1993) Isolation of virulence genes directing surface glycosyl-phosphatidylinositol synthesis by functional complementation of Leishmania. Proc Natl Acad Sci USA 90: 8609–8613.
31. King DL, Turco SJ (1988) A ricin agglutinin-resistant clone of Leishmania donovani deficient in lipophosphoglycan. Mol Biochem Parasitol 28: 285–294.
32. Boulanger N, Lowenberger C, Volf P, Ursic R, Sigutova L, et al. (2004) Characterization of a defensin from the sand fly Phlebotomus duboscqi induced by challenge with bacteria or the protozoan parasite Leishmania major. Infect Immun 72: 7140–7146.
33. Pimenta PFP, Modi GB, Pereira ST, Shahabuddin M, Sacks DL (1997) A novel role for the peritrophic matrix in protecting Leishmania from the hydrolytic activities of the sand fly midgut. Parasitology 115: 359–369.
34. Shakarian AM, Dwyer DM (2000) Structurally conserved soluble acid phosphatases are synthesized and released by Leishmania major promastigotes. Exp Parasitol 95: 79–84.
35. Ramalho-Ortigao M, Jochim RC, Anderson JM, Lawyer PG, Pham VM, et al. (2007) Exploring the midgut transcriptome of Phlebotomus papatasi: comparative analysis of expression profiles of sugar-fed, blood-fed and Leishmania major-infected sandflies. BMC Genomics 8: 300.
36. Jochim RC, Teixeira CR, Laughinghouse A, Mu JB, Oliveira F, et al. (2008) The midgut transcriptome of Lutzomyia longipalpis: comparative analysis of cDNA libraries from sugar-fed, blood-fed, post-digested and Leishmania infantum chagasi-infected sand flies. BMC Genomics 9: 15.

Research article

The midgut transcriptome of *Phlebotomus* (*Larroussius*) *perniciosus*, a vector of *Leishmania infantum*: comparison of sugar fed and blood fed sand flies

Anna Dostálová[1], Jan Votýpka[1], Amanda J. Favreau[2], Kent D. Barbian[2], Petr Volf[1], Jesus G. Valenzuela*[3] and Ryan C. Jochim*[3]

Address: [1]Department of Parasitology, Faculty of Science, Charles University in Prague, Vinicna 7, 128 44 Praha 2, Czech Republic, [2]Genomics Unit, Research Technologies Section, Rocky Mountain Laboratories, Hamilton, Montana 59840, USA, [3]Vector Molecular Biology Unit, Laboratory of Malaria and Vector Research, NIAID, NIH, Rockville, MD, 20852, USA

*Corresponding authors

AD: anna.svarovska@centrum.cz

JV: vapid@natur.cuni.cz

AJF: favrea2@mmc.org

KDB: kbarbian@niaid.nih.gov

PV: volf@cesnet.cz

JGV: jvalenzuela@niaid.nih.gov

RCJ: rjochim@niaid.nih.gov

Abstract

Background: Parasite-vector interactions are fundamental in the transmission of vector-borne diseases such as leishmaniasis. *Leishmania* development in the vector sand fly is confined to the digestive tract, where sand fly midgut molecules interact with the parasites. In this work we sequenced and analyzed two midgut-specific cDNA libraries from sugar fed and blood fed female *Phlebotomus perniciosus* and compared the transcript expression profiles.

Results: A total of 4111 high quality sequences were obtained from the two libraries and assembled into 370 contigs and 1085 singletons. Molecules with putative roles in blood meal digestion, peritrophic matrix formation, immunity and response to oxidative stress were identified, including proteins that were not previously reported in sand flies. These molecules were evaluated relative to other published sand fly transcripts. Comparative analysis of the two libraries revealed transcripts differentially expressed in response to blood feeding. Molecules up regulated by blood feeding include a putative peritrophin (*PperPer1*), two chymotrypsin-like proteins (*PperChym1* and *PperChym2*), a putative trypsin (*PperTryp3*) and four putative microvillar proteins (*PperMVP1*, *2*, *4* and *5*). Additionally, several transcripts were more abundant in the sugar fed midgut, such as two putative trypsins (*PperTryp1* and *PperTryp2*), a chymotrypsin (*PperChym3*) and a microvillar protein (*PperMVP3*). We performed a detailed temporal expression profile analysis of the putative trypsin transcripts using qPCR and confirmed the expression of blood-induced and blood-repressed trypsins. Trypsin expression was measured in *Leishmania infantum*-infected and uninfected sand flies, which identified the *L. infantum*-induced down regulation of *PperTryp3* at 24 hours post-blood meal.

Conclusion: This midgut tissue-specific transcriptome provides insight into the molecules expressed in the midgut of *P. perniciosus*, an important vector of visceral leishmaniasis in the Old World. Through the comparative analysis of the libraries we identified molecules differentially expressed during blood meal digestion. Additionally, this study provides a detailed comparison to transcripts of other sand flies. Moreover, our analysis of putative trypsins demonstrated that *L. infantum* infection can reduce the transcript abundance of trypsin *PperTryp3* in the midgut of *P. perniciosus*.

Background

Leishmaniases are a group of vector-borne diseases caused by parasitic protozoa of the genus *Leishmania*. *Leishmania infantum* (syn. *L. chagasi*) is the main etiological agent of visceral leishmaniasis, the most deadly form of the disease. The lack of a human vaccine, increasing resistance to the currently used drugs and their serious side effects urge the need for research of visceral leishmaniasis. In the western and central part of the Mediterranean basin, the major vector of *L. infantum* is *Phlebotomus perniciosus* [1].

Leishmania amastigotes are ingested with the blood meal upon female sand fly feeding on the mammalian host. After a series of morphological changes, propagation and migration of the parasites to the anterior part of the midgut, the infection is transmitted to another host during the next blood feeding. In hematophagous arthropods, blood feeding induces a number of processes including digestion, metabolism, diuresis and egg development. Unlike many other arthropod-borne infections, e.g., *Plasmodium* in mosquitoes, *Leishmania* complete their whole developmental cycle within the midgut of the sand fly. Several natural barriers to *Leishmania* development in the midgut have been described including the secreted proteolytic enzymes, the peritrophic matrix surrounding the ingested blood meal and the necessity to bind to the midgut epithelium (reviewed by [2]). Thus, the midgut is the primary organ where interactions between the vector sand fly and the parasite occur and it represents a key target for interruption of *Leishmania* transmission.

While the genome sequences of several *Leishmania* species, including *L. infantum*, have been published [3] and molecular studies abound, molecular data on sand flies are limited. An analysis of expressed sequence tags (ESTs) from the whole *Lutzomyia longipalpis* sand fly and salivary gland transcriptomes of several sand fly species have been published (reviewed by [4]). With regard to *Leishmania* development in the midgut, particularly midgut-specific transcriptomic analyses, studies of *L. longipalpis* and *Phlebotomus papatasi* [5-7] have brought important insights into the repertoire of molecules expressed in the midgut. Several midgut proteins from these two species were functionally characterized [8] and shown to impact *Leishmania* development [9,10].

In this study, we have generated and sequenced two cDNA libraries from the midgut tissue of *P. perniciosus* and analysed sequences present both before and after blood feeding. Furthermore, we provide phylogenetic analysis and comparison with the midgut molecules described in *L. longipalpis* and *P. papatasi*. Comparison of these three species is especially valuable with regard to *Leishmania* transmission. *P. papatasi* is the principal vector of cutaneous leishmaniasis caused by *Leishmania major* in the Old World [1]. It is refractory to the development of other species of *Leishmania* [11]. *Lutzomyia longipalpis* is the vector of *L. infantum* (*chagasi*) in Latin America and is considered a permissive vector due to full development of various *Leishmania* species in laboratory infections [1,12]. While being phylogenetically closer to *P. papatasi*, in some aspects *P. perniciosus* resembles *L. longipalpis*. First, it is a natural vector of *L. infantum*. Second, it is also permissive to the development of other *Leishmania* species [13]. Therefore, the present study provides a valuable database for identification of vector molecules that affect the vectorial competence of sand flies.

Results and Discussion

In order to gain insight into the spectrum of molecules present in the *P. perniciosus* midgut, two cDNA libraries from this organ were constructed, sequenced and analysed. The first library was constructed from a pool of midguts from sand flies allowed to feed on sucrose solution (sugar fed). For the construction of the second library (blood fed), midguts from sand flies 4-6h, 24h, 2, 3 and 4 days after blood feeding were pooled. These time points cover the course of blood digestion, allowing us to identify molecules transcribed in response to blood feeding. In total, 4511 clones were sequenced and 91% of the sequences were of high quality and included in subsequent analyses. Analysis was performed on 2049 and 2062 sequences for the sugar fed and blood fed libraries, respectively. These sequences were deposited in the NCBI dbEST database under accession numbers GW815603-GW820028. The comparable number of high quality sequences in each library allows for a better comparison of sequence abundance of specific molecules of interest in the libraries. The bioinformatic analyses of the sequences were performed using the dCAS cDNA annotation software [14]. Sequences were clustered together based on sequence homology and produced 207 and 163 contigs and 712 and 553 singletons

(cluster with only one sequence) for the sugar fed and blood fed libraries, respectively. The average sequence per contig ratio was higher in the blood fed library (9.26) than in the sugar fed one (6.46), attributed to the strong induction of certain sequences after blood feeding (such as sequences coding for putative microvillar proteins, proteolytic enzymes and peritrophins, as discussed later). Combining the two libraries produced 370 contigs, 1085 singletons and an average ratio of 8.18 sequences per contig. Most of the clusters (890) had a significant (E<10E-5) BLASTX match to the NCBI non-redundant protein database. However, 565 clusters, mostly singletons, produced no match or low homology and these clusters likely represent transcript coding for novel proteins or potential non-coding regions. Clusters were assigned to general functional classes using the best match BLAST results of the KOG database as a guideline. The overall distribution of clusters in these functional classes in the two libraries is shown in Figure 1. The distribution illustrates the abundance of microvillar proteins and proteins involved in amino acid transport and metabolism (a category including proteolytic enzymes) after blood feeding.

The following paragraphs give a detailed description of the most abundant sequences identified in the libraries and sequences that are of interest with respect to the midgut physiology and *Leishmania* life cycle. The sequences, their putative functions and distribution in the two libraries are listed in Table 1. Table 2 shows the best matches the sequences produced when compared to the NCBI non-redundant protein database using BLASTp.

Trypsins

Proteolytic enzymes were among the most abundant sequences detected in the libraries. Three putative trypsins were identified. *PperTryp1* (EZ933288, cluster 46) was one of the most abundant transcripts overall, strongly overrepresented in the sugar fed library (513 of 533 sequences). The putative protein has a predicted molecular weight of 27.6 kDa after cleavage of the signal peptide and a pI of 5.41. *PperTryp2* (EZ933289, cluster 16) is a less abundant (10 sequences) putative trypsin that was only detected in the sugar fed library. The putative mature protein has a predicted molecular weight of 26.9 kDa and a high pI of 8.83 (similar to a putative *P. papatasi* trypsin, PpTryp3). Sequences coding for a third putative trypsin named *PperTryp3* (EZ933290, cluster 63, 5` truncated) originated from the blood fed midgut library. In blood fed midguts we also identified a few partial transcripts, coding for a putative variant of this protein (5 sequences represented by clone PPRGFL_P8_H08, GW817404, Cluster 61). This Cluster 61 variant shows 82% identity to PperTryp3 at the amino acid level. Multiple sequence alignment of the putative *P. perniciosus* trypsin molecules (Figure 2) shows that structural cysteines, the H/D/S catalytic triad and putative substrate specifying residues are well conserved. Both *PperTryp1* and *PperTryp2*, for which we obtained the full-length sequence of the transcripts, are pre-pro-peptides; having a predicted signal peptide and a putative pro-peptide cleavage site for activation of the mature protein.

In order to describe the expression dynamics of the identified putative trypsin molecules, we performed a qPCR analysis of the three transcripts before, and at several time points after, blood feeding. The results (Figure 3) correlate with the sequence abundance in the two libraries, proving the validity of the library comparison approach. In addition, the qPCR analysis provides a more detailed view of the trypsin expression after blood feeding. *PperTryp1*, the most abundant trypsin identified, was down regulated as soon as 6h after blood feeding and further suppressed 24h post-blood meal (about $1/50^{th}$ of pre-blood meal levels). Its expression returned to the pre-blood meal levels in the sand flies that had passed the remnants of blood meal. *PperTryp2* was detected in lower amounts (about $1/70^{th}$ of PperTryp1) and represents another trypsin down regulated by blood feeding, with a time course similar to that observed for *PperTryp1*. In contrast, the qPCR analysis confirmed *PperTryp3* as the main blood feeding-induced trypsin molecule. *PperTryp3* expression was already elevated after 6h and the highest quantity of the transcripts was observed 24h post-blood meal. *PperTryp3* returned to negligible amounts in sand flies that had finished blood digestion.

This study brings the first expression analysis of sand fly trypsins using precise quantification by the means of qPCR. The observed *P. perniciosus* trypsin expression profile is in accordance with the results of earlier studies of *P. papatasi* and *L. longipalpis* midgut trypsin abundance from data acquired by semi-quantitative end-point PCR [15,16] and comparison of transcript abundance in cDNA libraries before and after blood feeding [5,6]. In all the three species, one

or several trypsin transcripts are present in high abundance in sugar fed females while their quantities decrease after the intake of blood (*PperTryp1*, *PperTryp2*, *LlTryp2*, *PpTryp1*, *PpTryp2*). At the same time, the expression of other putative trypsins (*PperTryp3*, *LlTryp1* and *PpTryp4*) is induced upon blood feeding.

Phylogenetic analysis of the putative trypsins (Figure 4) shows that the sequences abundant before blood feeding share similarity and, together with LuloTryp4 (reported in similar numbers both from blood fed and sugar fed *L. longipalpis* sand flies), form a clade apart from the other sand fly trypsins that include PperTryp3. The predicted pre-pro-peptide structure and high abundance of transcripts in sugar fed sand flies together with the virtual absence of trypsin-like enzymatic activity in sugar fed sand fly midguts [17] suggests that PperTryp1, PperTryp2 and their respective homologues are associated with initial blood meal digestion as they can be quickly translated and processed following blood feeding.

The onslaught of proteolytic activity after the intake of blood is one of the barriers for *Leishmania* development in the midgut [2]. Sant`Anna et al. [10] have shown that suppression of the major blood meal-induced trypsin (LlTryp1) in *L. longipalpis* by the means of RNAi enhances the survival of *L. mexicana* in the midgut. Some of earlier studies have demonstrated the ability of *L. major* to suppress or delay the peak of trypsin activity in the midgut [18,19]. More recent transcriptomic studies [5,6] have shown modulation of trypsin-like transcript abundance in *P. papatasi* and *L. longipalpis* midgut in the presence of *Leishmania* parasites. qPCR was used to measure trypsin expression in infected sand flies to evaluate modulation of the *P. perniciosus* midgut trypsins by *L. infantum*. Our analysis showed that the amount of *PperTryp3*, the major blood meal-induced trypsin, is decreased in the presence of *L. infantum* (Figure 3). This difference was observed in sand flies 24 hours post-blood meal, which correlated with the peak of *PperTryp3* expression in uninfected blood fed sand flies. Our findings suggest the ability of *L. infantum* to suppress or delay the expression of the major blood-induced trypsin in *P. perniciosus* and identify this molecule as an interesting candidate for future studies.

Chymotrypsins

Chymotrypsin-like enzymes are another group of proteases found in abundance in the midgut of sand flies and mosquitoes. Five clusters coding for putative chymotrypsins were identified in the libraries and each cluster was 5` truncated. *PperChym1* (EZ933296, cluster 81) was the most abundant and was only detected in the blood fed midgut library (82 sequences). Similarly, the second most abundant chymotrypsin-like sequence, *PperChym2* (EZ933297, cluster 102) probably codes for a digestive enzyme up-regulated by blood feeding as 11 of the 13 sequences were found in the blood fed library. In contrast, *PperChym3* (EZ933298, cluster 1033) and *PperChym4* (EZ933299, cluster 816) sequences were only found in the sugar fed sand fly library. The expression pattern of these chymotrypsins is similar to the aforementioned trypsin molecules indicating that there may be early and late classes of serine proteases in sand flies, similar to what has been observed in mosquitoes[20] The least abundant putative chymotrypsin sequence *PperChym5* (EZ933300, cluster 710) was represented by one sequence in each of the libraries. The phylogenetic analysis of amino acid sequences (Figure 5) shows conservation in sequence homology of PperChym1-4 with putative *P. papatasi* and *L. longipalpis* midgut chymotrypsins. PperChym2 formed a subclade along with LuloChym2 and Ppchym3. PperChym5 is more distantly related to putative chymotrypsins described in other sand flies. The most similar sequence in the NCBI non-redundant protein database is a putative serine protease of *Culex quinquefasciatus* (XP_001845462.1, E= 1 e-59) and the best match in the Swissprot database is the white shrimp, *Litopenaeus vannamei,* Chymotrypsin BI (Q00871, E= 3 e-36) with a proven chymotrypsin catalytic activity. Also, a conserved serine residue at the substrate specifying site suggests a chymotrypsin-like specificity of PperChym5 enzyme (Figure 6). The H/D/S catalytic triad and cysteine residues are well conserved among all the putative *P. perniciosus* chymotrypsin sequences. Putative chymotrypsin transcript abundance has previously been shown to be altered by *Leishmania* infection in the midgut. *LuloChym1A* in *L. longipalpis* and *PpChym2* in *P. papatasi* were reported as underrepresented in the midgut in the presence of *L. infantum* and *L. major*, respectively[5,6].

Carboxypeptidases

A number of clusters coding for putative proteins with homology to carboxypeptidases were identified. Two putative metallo-carboxypeptidases of the M14 A/B subfamily were named *PperCpepA* and *PperCpepB*. *PperCpepA* (EZ966131, cluster 539, 5` truncated) shows similarity to carboxypeptidases A described from mosquitoes. Phylogenetically, PperCpepA clusters more distantly from the carboxypeptidases A described in the midgut of other phlebotomine species (Figure 7). Carboxypeptidase A, a zinc-metalloprotease, can hydrolyze aromatic and aliphatic side chains from the C-terminus. *PperCpepB* (EZ966132, cluster 217, 5` truncated) is similar to mosquito and sand fly midgut carboxypeptidase B. Carboxypeptidase B specifically hydrolyzes C-terminal arginine and lysine. PperCpepB possesses the conserved aspartate residue at the position responsible for this specific substrate recognition [21] (Figure 8). Due to the low number of sequences in this cluster a comparative analysis between the sugar fed and blood fed libraries was not possible; however, it is notable that five of the six sequences of *PperCpepB* were contributed by the blood fed library. *Anopheles gambiae* midgut carboxypeptidase B has been shown to be up-regulated by *Plasmodium* infection and antibodies against one of these enzymes, CPBAg1 (CAF28572) blocked parasite development in the mosquito midgut [22]. In *L. longipalpis*, one of the carboxypeptidases transcripts (LuloCpepA1) was underrepresented in a cDNA library from *L. infantum*-infected midgut as compared to uninfected sand flies [6].

Aminopeptidases

A partial transcript coding for a putative alanyl aminopeptidase was identified (cluster 126). The molecule, named *PperApeptN* (EZ966135), is similar to mosquito membrane aminopeptidases of the M1 family (aminopeptidase N). It was abundant in both libraries (13 and 7 sequences in blood fed and sugar fed midgut libraries, respectively). Membrane alanyl aminopeptidases were described in the midgut of many Dipteran species including mosquitoes, where they were identified as receptors for *Plasmodium* ookinetes and also *Bacillus thuringiensis* Cry toxin binding [23,24]. In the sand fly midgut, aminopeptidase activity was detected after blood feeding, mainly associated with the midgut wall (using leucine-p-nitroanilide LpNA as a substrate) [17]. This activity was reduced in *P. papatasi* and *P. langeroni* following infection with *L. major* [19].

Astacins

Two clusters coding for putative astacin-like zinc metalloproteases were identified in the libraries. The more abundant cluster, *PperAstacin1* (EZ966133, cluster 84), is predicted to encode a protein with a molecular weight of 27.0 kDa once secreted and pI of 5.05. It was present both in the sugar fed and blood fed libraries. The transcript of cluster 967 was named *PperAstacin2* (EZ966134) and the predicted translated product has a molecular weight of 26.5 and pI 6.00 after cleavage of the signal peptide. It was only detected in the sugar fed library. Phylogenetic analysis of other putative astacin sequences shows that PperAstacin1 is similar to astacin-like molecules previously described in *L. longipalpis* (LuloAstacin), *P. papatasi* and other Diptera. PperAstacin2 is most similar to a putative astacin from *A. gambiae* using BLASTp similarity search of the NCBI non-redundant protein database. However, in a phylogenetic analysis it branches away from all other Dipteran sequences (Figure 9A). Multiple sequence alignment (Figure 9B) shows the differences in amino acid sequences and illustrates the conservation of all residues likely responsible for zinc-binding and catalytic activity in the putative *P. perniciosus* astacins.

Microvillar proteins

The most abundant transcripts identified in the library were sequences coding for proteins with similarity to major insect allergen proteins. These insect-specific proteins containing insect-allergen domains (InterPro IPR010629) were first described as the major human allergens in the faeces of the cockroaches *Blatella germanica* and *Periplaneta americana* [25]. In butterflies of the Pieridae family, a novel family of proteins with multiple insect-allergen domains has evolved (nitrile-specifier protein family) to serve a role in detoxification of plant metabolites in the butterfly larvae food [26,27]. In mosquitoes, proteins with a single insect-allergen domain have been identified and termed G12 microvillar proteins. These molecules have been shown to be induced in the mosquito midgut after blood-feeding [28,29]. In *Aedes aegypti*, the G12 protein (AEG12, AAL05408.1) has been shown to be expressed only in the midgut after blood feeding and located on the microvillar membranes of the midgut epithelial cells [29]. The role of insect

allergen proteins, other than nitrile-specifier protein family in Pieridae, has not yet been characterized.

We identified five putative homologues of the insect-allergen proteins in the *P. perniciosus* libraries. These putative microvillar proteins (MVPs) possess a predicted signal peptide (where full-length sequences were obtained) and a single insect-allergen domain. *PperMVP1* (EZ933291) encodes a putative protein with a mature molecular weight of 21.4kDa and pI of 5.16. Derived from 681 sequences (cluster 45 and variants) *PperMVP1* was the most abundant transcript overall. Transcripts of *PperMVP1*, and three other *P. perniciosus* MVPs, were only found in the blood fed cDNA library. These other blood feeding induced MVPs include *PperMVP2* (EZ933292, cluster 40 and variants, sequence 5` truncated), *PperMVP4* (EZ933294, cluster 139, 5` truncated) and *PperMVP5* (EZ933295, cluster 52 and variants; 22 kDa, pI 4.77). PperMVP4 amino acid sequence, although truncated, contains predicted glycosylation sites (3 N-glycosylations). The only MVP transcript overrepresented in the sugar fed library is *PperMVP3* (EZ933293, cluster 52 and variants, 5` truncated). As demonstrated in the phylogenetic analysis (Figure 10A), the identified sequences show high similarity to the respective five MVPs previously identified in the midgut of *L. longipalpis*. Homologues of four of these proteins are also known in *P. papatasi* (PpMVP1-4; Figure 10A). Interestingly, no sequence with high similarity to PperMVP5 was found in the midgut of *P. papatasi*. The phylogenetic tree also shows that PperMVP3 and its putative orthologues LuloMVP3 and PpMVP3 clade away from all the other sand fly and mosquito MVPs. This is in accordance with the fact that all the three seem to be down-regulated be blood feeding unlike other MVPs. Multiple sequence alignment (Figure 10B) shows that the five putative *P. perniciosus* MVPs share little sequence homology suggesting that these molecules may have different functions altogether.

Antimicrobial molecules

Several transcripts encode proteins putatively involved in the immune response of the sand fly midgut. Two clusters coding for putative peptidoglycan recognition proteins (PGRPs) were identified. PGRPs play central and diverse roles in activating insect immune reactions including the melanization cascade, phagocytosis, and signal transduction pathways for production of antibacterial peptides. *PperPGRPLB* (EZ617708, cluster 274) is similar to a PGRP previously reported from *L. longipalpis* (*LuloPGRP*) and *P. papatasi* (*PpPGRP*) midgut. Although the transcript appears to be 5` truncated, based on homology to other full length sand fly transcripts we predict it to be a protein of about 22kDa. A homologous (but secreted) protein PGRPLB in *Drosophila melanogaster* (AAN13505), mainly expressed in the midgut, was shown to regulate the Imd signalling pathway controlling the immune response against Gram-negative bacteria [30]. An *A. gambiae* PGRPLB homologue (EAA01800) is induced by both bacterial and *Plasmodium* infections [31]. *PperPGRPLC* **(**EZ617707, cluster 168, transcript 5` truncated**)** encodes a putative protein with similarity to insect PGRPLC proteins that act as membrane-localized peptidoglycan receptors activating the Imd pathway. PGRPLC-like molecules have not been previously reported in sand flies. In searching the midgut transcriptomes of sand flies, partial transcripts with homology to *PperPGRPLC* were found in *P. papatasi* (ES347179) and *L. longipalpis* (AM098991). In *A. gambiae*, PGRPLC (AGAP005203) signalling controls the size of symbiotic bacteria populations, intestinal bacterial infections and *Plasmodium* infections [32].

In addition to PGRPs, a transcript encoding a putative gram-negative bacteria binding protein was identified and named *PperGNBP* (EZ617706, cluster 301, transcript 5` truncated). Gram-negative bacteria-binding proteins serve as pattern recognition receptors binding to pathogen-associated beta-1,3-glucans in insects and they have been shown to play a role in mosquito defence against bacteria and *Plasmodium* infection [33].

It is likely that the identified sand fly pattern recognition proteins are involved in protection against bacteria in the midgut; however, similar to the mosquito homologues, they might also have an impact on *Leishmania* infection. Interestingly, Kumar et al. [34] have recently described a novel secreted peroxidase/dual oxidase system in *A. gambiae* midgut that catalyses cross-linking of a dityrosine network on the luminal surface of the epithelial cells. This network decreases the permeability of the mucus layer to immune elicitors and thus prevents induction of PGRPs, nitric oxide synthase and other immune responsive genes. Silencing of the peroxidase/dual oxidase system causes a drastic reduction in *Plasmodium* infection in the

midgut [34]. We have not identified any homologues of the components of the peroxidase/dual oxidase system in *P. perniciosus* midgut in this analysis. Further studies are needed to see whether a dityrosine barrier is formed in the sand fly midgut. In our analysis, we did not detect any antimicrobial peptide transcripts in the midgut of *P. perniciosus*, although a defensin was previously characterized in *Phlebotomus duboscqi* midgut [35] and defensin transcripts were also reported from the midgut of *L. longipalpis* [6,7].

Oxidative stress molecules

A number of transcripts were identified coding for putative antioxidant enzymes. In hematophagous insects, blood meal-derived free heme is a strong pro-oxidant and can tax the midgut antioxidant system. In addition to their protective role, redox-related molecules were shown to regulate midgut epithelial immunity and impact the outcome of bacterial and parasitic infections in mosquitoes [34,36]

Glutathione S-tranferases (GSTs) and peroxiredoxin

Putative components of the glutathione-thioredoxin system, the central redox homeostasis maintaining pathway in insects, were found [37,38]. Several transcripts coding for putative glutathione-S-transferases (GSTs), enzymes catalyzing substrate detoxification by a thiol tripeptide glutathione, were identified in both libraries. *PperGST1* (EZ617709, cluster 163) encodes a putative intracellular GST of the Sigma subfamily. The putative protein is predicted to be 23.2 kDa and have a pI of 5.00. It is nearly identical to putative Sigma GSTs described from the midgut of *P. papatasi* (ABV44736, 98% identity) and *L. longipalpis* (LuloGST1, 97% identity). *PperGST2* (EZ617710, cluster 463, 5` truncated) encodes a putative transmembrane protein that has homology to microsomal GSTs of the MAPEG super family. Homologues of *PperGST2* were found in the EST databases of *P. papatasi* (FK811479) and *L. longipalpis* (EW990920). *PperGST3* (EZ617711, cluster 1322, 5` truncated) shares similarity with other Dipteran GSTs of the Theta class, a class not previously reported in sand flies. A homolog of PperGST3 was identified by searching the *L. longipalpis* whole fly cDNA library (AM099640) [39].Unlike the midgut transcriptomes of *P. papatasi* and *L. longipalpis*, we have not found any GSTs of the Delta/Epsilon class, which may be due to the overall low abundance of the GST transcripts in sand fly midguts. Mosquito GSTs play an important role in as antioxidants and knock-down of GSTs of the Theta family has been reported to impact *Plasmodium* infections in *A. gambiae* and *A. stephensi*, although the effect varies with different parasite-vector combinations [40].

A putative peroxiredoxin (or thioredoxin-dependent peroxidase), product of cluster 729, was identified and named *PperPRX* (HM119220). *PperPRX* encodes a putative intracellular protein of 16.7 kDa and a pI of 7.1 containing a peroxiredoxin PRX5-like subfamily domain. Salp25D, a peroxiredoxin in the tick *Ixodes scapularis*, has been shown to facilitate the acquisition of *Borrelia* from an infected host by detoxifying reactive oxygen species at the vector-pathogen-host interface [41]. Midgut-specific Salp25D, while not significantly aiding the establishment of *Borrelia*, does have a slight protective effect. It is possible that sand fly peroxiredoxins, by detoxifying OH radicals, could have a similar protective effect on *Leishmania* parasites.

Catalase and superoxide dismutases (SODs)

Transcripts coding for putative enzymes of the superoxide dismutase (SOD)/catalase system were also identified. *PperCat* (EZ617712, cluster 852) encodes a putative intracellular protein (molecular weight 57.7 kDa and pI 9.17) containing a catalase domain. It shares high similarity with a putative catalase molecule described in *L. longipalpis* midgut (*LuloCat*) and a similar sequence was also found in the *P. papatasi* midgut cDNA library (ES351062). Catalases are hydrogen peroxide detoxifying enzymes and for an *A. gambiae* homologue (AGAP004904), expression is induced in the midgut after blood feeding in response to oxidative stress [42]. In the *P. perniciosus* midgut, the transcript was only found in the sugar fed library (9 sequences) and thus appears to be down regulated by blood feeding. A similar phenomenon could not be observed in *L. longipalpis* due to the low number of catalase sequences found (where one *LuloCat* transcript was found in the blood fed and one in the post-blood fed library infected with *L. chagasi*). The significance of *PperCat* down regulation by blood feeding remains unclear and post-transcriptional regulation cannot be excluded.

Two clusters with products containing copper-zinc superoxide dismutase (Cu-Zn SOD) domains were identified. *PperSOD1* (EZ617713, cluster 892) encodes a protein similar to a putative secreted Cu-Zn SOD from the midgut of *L. longipalpis* (LuloSOD). Despite the transcript being 5` truncated, based on homology to the *Lutzomyia* and mosquito molecules we predict the protein possesses a signal peptide and performs a similar function to LuloSOD. The molecule may be secreted or, given that the sequence contains a putative GPI-anchor site, GPI-anchored to the plasma membrane of the midgut cells. *PperSOD2* (EZ617714, cluster 1166) encodes a putative intracellular protein (15.3kDa, pI 6.3) similar to putative mosquito SODs. Intracellular SODs have not been previously described in the sand fly midgut, but when searched for homologous sequences, we found sequences coding for highly similar proteins in both *L. longipalpis* and *P. papatasi* midgut cDNA libraries (EW987718 and ES348811, respectively). Phylogenetic analysis of mosquito and sand fly sequences (Figure 11) shows that extracellular and intracellular Cu-Zn SODs form two distinct clades suggesting the two forms of the enzymes evolved prior to speciation of the two groups of organisms.

Xanthine dehydrogenase

Cluster 373 (EZ617715, *PperXDH*) is a partial transcript coding for a product with high similarity to the C-terminal portion of *L. longipalpis* xanthine dehydrogenase (XDH) (CAP08999.1). XDHs catalyze the oxidation of xanthine to urate, the main product of nitrogen metabolism, which has antioxidant properties in insects. The XDH molecule has been shown to be up regulated after blood feeding in *L. longipalpis*. Silencing of *L. longipalpis* XDH resulted in a reduction in urate production and a reduced life span of both sugar fed and blood fed sand flies [43]. These results suggest that xanthine dehydrogenases are indeed involved in preventing oxidative damage by producing the antioxidant urate in sand flies.

Ferritin

Two molecules were identified with homology to ferritins described from other insect species including the sand flies *P. papatasi* and *L. longipalpis*. Similar to most insect ferritins, and unlike most vertebrate and plant ferritins, these molecules are likely secreted. *PperFLC* (EZ617716, cluster 88) encodes a putative protein with homology to ferritin light-chain subunit and a molecular weight of 24.3kDa and pI 6.68, once secreted. *PperFHC* (EZ617717, cluster 332) encodes a putative ferritin heavy-chain molecule. Although 5` truncated, *PperFCH* has high homology to *P. papatasi* and other species ferritin molecules with signal peptides and is likely similarly secreted. The transcripts originated in similar numbers from the blood fed and sugar fed libraries. Apart from their role in iron metabolism, these sand fly midgut ferritins may also play a role in preventing oxidative damage by sequestering large quantities of free iron from the digested blood meal as was suggested for other blood-feeding insects [44].

Peritrophic matrix proteins and chitinase

The peritrophic matrix (PM) is an extracellular chitin-containing matrix that is formed in the sand fly midgut after blood feeding that surrounds the ingested blood. Clusters coding for products with similarity to PM proteins described in other blood feeding Dipterans were identified. Three clusters coding for putative peritrophins were detected, originating in higher numbers from the blood fed library. These proteins share homology with molecules previously identified in other sand fly species and contain chitin binding domains (CBDs). *PperPer1* (EZ933302, cluster 97) encodes a protein similar to putative peritrophins with four CBDs previously described from the midgut of *L. longipalpis* (LuloPer1) and *P. papatasi* (PpPer1). *PperPer1* is represented by 94 sequences found only in the blood fed library. Although *PperPer1* ESTs appears to be incomplete at the 5` end and missing the first 13 N-terminal amino acids, based on homology, we predict it to be a secreted molecule of mature molecular weight of 28.2 kDa and pI 4.68. *PperPer2* (HM119221, cluster 330) encodes a putative peritrophin with similarity to a *L. longipalpis* protein LuloPer2 containing one CBD. A third putative peritrophin, *PperPer3* (HM119222, cluster 156), shows similarity to a *P. papatasi* peritrophin PpPer3 and contains two putative CBDs (one partial CBD sequenced due to a 5` truncation). Phylogenetic analysis (Figure 12) of the CBDs from *P. perniciosus*, *P. papatasi* and *L. longipalpis* illustrates a high degree of conservation of the Peritrophin1 arrangement. The four peritrophin domains share respective homology in all the three sand fly species.

A cluster coding for a putative chitinase was identified in the blood fed library. The molecule, named *PperChit* (EZ933285, cluster 124) encodes a putative protein containing a CBD and a mucin-like domain and shares high similarity to *P. papatasi* PpChit1 and *L. longipalpis* (LlChit1)

chitinases. These midgut-specific, blood meal-induced enzymes have been shown to account for chytinolytic activity in the sand fly midgut and have been implicated in the release of *Leishmania* parasites from the endoperitrophic space [8,45,46].

Along with peritrophins and chitinases, non-chitin binding peritrophic matrix proteins have also been described from bloodsucking insects. Dinglasan et al. [47] performed a proteomic analysis of *A. gambiae* peritrophic matrix and identified a number of non-chitin-binding proteins including proteolytic enzymes and novel classes of PM proteins with unknown function. We identified clusters with homology to these proteins in the *P. perniciosus* libraries. The product of cluster 379 (EZ933287) is highly similar to *A. gambiae* PM protein AGAP006398. The transcript contains a potential N-glycosylation site and 3 DM9 repeats (repeats of unknown function found in a number of arthropod proteins). A homolog was found in the *P. papatasi* midgut library (31.5 kDa midgut protein ABV44721). The product of cluster 358 (EZ933286, 5` truncated) showed a significant similarity (BLASTp match 4e-18) to an *A. gambiae* PM protein AGAP000570. The sequences originated from both blood fed (3 sequences) and sugar fed (10 sequences) libraries. Homologous proteins were also found in the midgut of other phlebotomines (ABV44744 in *P. papatasi* and ABV60298 in *L. longipalpis*) and also salivary glands of *P. duboscqi* (ABI20163). These proteins contain no described conserved domains. Based on homology to the *A. gambiae* protein found in the PM, we speculate that the product of cluster 358 may also be involved in the PM formation in sand flies. On the other hand, its expression in the sugar fed midgut as well as the presence of homologs in the salivary glands may suggest a different function for this protein, such as regulating the haemostatic response.

The structure of the mosquito and sand fly peritrophic matrix is complex and rearranges during the course of blood digestion [48]. The two putative peritrophins with multiple CBDs (PperPer1 and PperPer3) are likely to have a role in cross-linking the chitin fibrils of the peritrophic matrix. In addition to chitin binding, mosquito proteins with CBDs have also been described to bind heme and have a role in its sequestration during blood digestion [49]. Also, the glycosylation of the PM proteins can be of great importance for the PM structure and function. Two of the putative peritrophins, PperPer2 and PperPer3, and the putative chitinase, PperChit, contain mucin-like (Pro-Ser/Thr rich) domains. Glycosylation of these domains can influence the selectiveness of the PM pores, account for water retention within the PM and also protect the molecules from degradation by proteolytic enzymes. Furthermore, the degradation of putatively aglycosylated PM proteins (like PperPer1 and the product of cluster 358) by temporally secreted digestive proteases may play a role in the changes in the PM thickness and structure.

Transcripts differentially expressed after blood feeding

In order to identify changes in expression of midgut proteins induced by blood feeding, we compared the abundance of transcripts in the sugar fed and the blood fed libraries using chi-square statistical analysis. We found several transcripts that were significantly more abundant after blood feeding and several that were underrepresented in the blood fed library (see tables 3 and 4). As expected, we observed some transcripts putatively involved in blood digestion and peritrophic matrix formation more abundant after blood feeding. These included the putative peritrophin with four chitin-binding domains, *PperPer1*. Similar to the putative orthologues in *P. papatasi* and *L. longipalpis* (*PpPer1*, EU031912 and *LuloPer1*, EU124588), *PperPer1* was only detected in the blood fed midgut library. With regard to peritrophin sequence abundance, it is interesting to note that we did not detect any peritrophin sequence highly represented before blood feeding. This is in contrast with the situation described in *P. papatasi*, where high numbers of a putative peritrophin with one chitin binding domain (*PpPer2*, EU047543) were detected in sugar fed midguts. In this respect, the observed profile is more similar to peritrophins in the midgut of a more distantly related species *L. longipalpis*.

Transcripts coding for proteolytic enzymes, namely the chymotrypsins *PperChym1* and *PperChym2* and the trypsin *PperTryp3*, were also found more abundant in the blood fed library and thus likely represent digestive enzymes induced by the intake of blood. On the other hand, three other putative proteases, *PperTryp1*, *PperTryp2* and *PperChym3* were significantly less abundant in the blood fed library. We speculate that these molecules may be post-transcriptionally regulated digestive enzymes that are stored in the midgut prior to blood feeding.

The most striking differences in sequence abundance before and after blood feeding were observed for the microvillar proteins. Four of the five identified putative microvillar proteins (*PperMVP1*, *2*, *4* and *5*) were only detected in the blood fed library and in high abundance. This indicates a strong up-regulation of these proteins after the intake of blood. In contrast, *PperMVP3* was overrepresented in the sugar fed library, suggesting a different role for this protein. The observed microvillar proteins EST distributions are in accordance with what was described in both *P. papatasi* and *L. longipalpis*, where the *PperMVP3* orthologues, *LuloMVP3* and *PpMVP3* were highly represented before blood feeding unlike all other microvillar proteins. The conservation of these proteins and their pattern of expression in the midgut of the three sand fly species indicate their important, yet uncharacterized, roles in the midgut physiology.

The list of sequences overabundant in the blood fed library also includes a putative protein (cluster 79, HQ015441) similar to a putative cockroach allergen MPA2 and several uncharacterized mosquito proteins. The presence of a lipid-binding ML domain in the translated sequence of cluster 79 (Interpro IPR003172) may suggest a role of lipid recognition. In accordance with their putative function in carbohydrate digestion, putative glycoside hydrolases (cluster 174, HQ015444 and cluster 183, HQ015443) were found overrepresented in the sugar fed library. Interestingly, we also found a putative catalase sequence (PperCat) overabundant in the sugar fed midgut. Similarly, the significance of the higher abundance of a putative 40S ribosomal protein SA (cluster 652, HQ015442) and an unknown, probably non-coding, sequence (cluster 249, represented by the clone GW817178) in the sugar fed midgut remains unknown.

Conclusion

P. perniciosus is a medically important vector of canine and human visceral leishmaniasis in the Old World. To date, the only molecular data available for this species have been the salivary gland transcripts [50]. This study is the first report on molecules present in the midgut of *P. perniciosus*. As development of *Leishmania* in the vector sand fly is restricted to the digestive tract, the midgut is the primary organ where interactions with *Leishmania* take place. By sequencing and analyzing transcripts present before and after blood feeding, we have provided a catalogue of putative proteins potentially involved in feeding and blood digestion. All the generated ESTs were deposited in the NCBI dbEST database, making them available to scientific communities for further research. Selected molecules of interest were manually annotated and the nucleotide and putative protein sequences submitted to GenBank.

We have identified a variety of molecules, including putative proteins that have not been previously described in the sand fly midgut. Among the putative proteases, these include a putative astacin (PperAstacin2) and a putative chymotrypsin (PperChym5). We also found molecules potentially involved in pathogen recognition such as the gram-negative bacteria binding protein (PperGNBP) and the putative peptidoglycan receptor (PperPGRPLC). Novel putative antioxidant enzymes were also identified including an intracellular superoxide dismutase (PperSOD2) and putative microsomal and Theta class glutathione S-transferases (PperGST2 and PperGST3). In addition, we describe homologs of mosquito peritrophic matrix proteins.

Constructing libraries from sand fly females before and after the intake of blood allowed for the identification of molecules differentially expressed in response to blood feeding. By comparing our findings with the midgut transcriptome analyses of two other sand fly species, *L. longipalpis* and *P. papatasi*, we identified several features shared by the two permissive vectors, *P. perniciosus* and *L. longipalpis*. These include the absence of a significant number of peritrophin sequences before blood feeding, while in *P. papatasi*, a peritrophin with one chitin binding domain was abundant in sugar fed midguts.

Altogether, this study contributes to our knowledge of the molecular background of events that occur in the sand fly midgut. It provides a valuable platform for functional studies of selected molecules relevant in the transmission of *Leishmania*. These may represent targets for use as novel vector-based transmission-blocking vaccines to control this neglected disease.

Methods

Sand fly maintenance and dissection

The colony of *Phlebotomus perniciosus* (originally from Spain) was maintained in the insectary of Charles University in Prague as described previously [51]. Adults were kept at 26 °C and fed on 50% sucrose ad libitum. Ten midguts from 3-5 days old sugar fed only females were dissected for the sugar fed library construction. Females were fed on an anaesthetised Balb/C mouse and two midguts containing blood were dissected at each of the following time points: 4-6h, 24h, 2 days, 3 days and 4 days post-blood meal. These samples were pooled for the construction of the blood fed library. For the qPCR experiment, females were fed through a chick skin on heat-inactivated rabbit blood containing *L. infantum* infected macrophages (or uninfected macrophages in the control group). Midguts from ten infected and ten uninfected (control) females were dissected 6h, 24h, 72h and 10 days after blood feeding and stored individually in 10µl of RNAlater (Ambion) as well as 10 individual midguts from sand flies before blood feeding (sugar fed). Presence of promastigotes in midguts was confirmed 72h after blood feeding. On day 10 after blood feeding only sand flies with late-stage infections (with parasites on the stomodeal valve) were used.

cDNA library construction and sequencing

Messenger RNA was purified from midguts stored in RNAlater (Ambion) using the MicroFastTrack mRNA isolation kit (Invitrogen). PCR-based cDNA libraries were prepared following the instructions for the SMART cDNA library construction kit (Clontech). Each cDNA library was then fractionated into three sets containing small, medium and large fragments using columns provided by the manufacturer. Concentrated cDNA was ligated into a lambda TriplEx2 vector (Clontech). The resulting ligation reaction was packed using the Gigapack III Gold (Stratagene). The libraries thus obtained were plated by infecting log-phase XL1-blue cells (Clontech). Phage plaques lacking β-galactosidase activity were picked using sterile wooden sticks and placed into 75µl of water. Amplification of the cDNA was performed using Faststart Mix (Roche), 3µl template and primers PT2F1 (5'-AAG TAC TCT AGC AAT TGT GAG C-3') and PT2R1 (5'-CTC TTC GCT ATT ACG CCA GCT G-3'). The PCR conditions were 1 hold of 75 °C for 3 min, 1 hold of 94 °C for 4 min, 33 cycles of 94 °C for 1 min, 49 °C for 1 min, and 72 °C for 7 min, 1 hold of 75 °C for 7 min. The amplification product was cleaned with three washes with ultra pure water using ExcelaPure plates (EdgeBio) resuspended in 30µl of water. Sequencing was performed at the Rocky Mountain Laboratories Genomics Unit as described previously [52]. The template was combined with primer PT2F3 (5'- TCT CGG GAA GCG CGC CAT TGT-3') in an ABI 96-well Optical Reaction Plate (P/N 4306737) following the manufacturers recommendations. Sequencing reactions were setup as recommended by Applied Biosystems BigDye Terminator v3.1 Cycle Sequencing Kit by adding 1 µL ABI BigDye Terminator Ready Reaction Mix v3.1 (P/N 4336921), 1.5 µL 5× ABI Sequencing Buffer (P/N 4336699), and 3.5 µL of water for a final volume of 10 µL. Cycle sequencing was performed at 96 °C for 10 s, 50 °C for 5 s, 60 °C for 4 min for 27 cycles on either a Bio-Rad Tetrad 2 (Bio-Rad Laboratories, Hercules, CA) or ABI 9700 (Applied Biosystems, Inc., Foster City, CA) thermal cycler. Fluorescently labelled extension products were purified following Applied Biosystems BigDye XTerminator Purification protocol and then processed on an ABI 3730xL DNA Analyzer (Applied Biosystems, Inc., Foster City, CA).

Bioinformatic analysis

The bioinformatic analyses of the sequences were performed using the dCAS 1.4 cDNA annotation software [14]. Briefly, primer, vector and low quality sequences were removed at the 5' and 3' ends of each sequence using Cap3 and Phred software [53-55]. Sequences from both libraries were grouped together and aligned to generate clusters based on 95% identity over 100 nucleotides. Three frame translated consensus sequences were supplied to the appropriate BLAST algorithm [56] for comparison to the contents of the NCBI non-redundant protein database, the Gene Ontology database [57], the KOG conserved domain database [58], Simple Modular Architecture Tool (SMART) [59], Protein Family Database (Pfam) [60], rRNA subset database and Mitochondrial and Plasmid Sequences database (MIT-PLA) available from NCBI. The predicted presence of a signal secretion peptide or transmembrane domains was determined using the SignalP and TMHMM programs respectively [61,62]. N- and O-glycosylation site prediction was performed for selected sequences using NetNGlyc 1.0 and NetOGlyc 3.1 software (http://www.cbs.dtu.dk/services/NetNGlyc/) [63]. Numbers of sequences in the sugar fed and the blood fed library were compared using χ^2 statistical analysis. Clusters

with significantly unequal distribution of the clone sequences (P<0.05 and expected frequency in each of the libraries > 4) were identified as over- or underrepresented after blood feeding. Selected sequences were aligned using Clustal X 2.0 [64] and manually refined in BioEdit 7.0 sequence-editing software. For phylogenetic analyses of amino acid sequences, best substitution matrix was determined for each alignment by ProtTest software, version 2.0 [65]. This matrix was then used by TREE-PUZZLE 5.2 [66] to reconstruct maximum likelihood phylogenetic trees from the protein alignments using quartet puzzling with 1000 puzzling steps. Resulting trees were visualized in MEGA 4 [67].

Macrophage infection

Leishmania infantum (MCAN/PT/05/IMT 373) parasites were cultured at 23 °C in RPMI medium (Sigma) containing 10% heat-inactivated foetal calf serum (FCS, Gibco), 50 ug/ml gentamicin, 1x BME vitamins (Sigma) and 1% human urine. Mouse macrophage line J774 was cultured at 37 °C, 5% CO_2 in RPMI medium containing 10% FCS, 2mM alanyl-glutamine and penicillin (200 IU/ml). Macrophages were infected with stationary-phase *L. infantum* parasites at 1:10 macrophage:parasite ratio. After 24h of co-cultivation at 37 °C, 5% CO_2, non-internalized parasites were removed by washing 3x in the culture media and infected macrophages cultivated overnight. Macrophages were confirmed to contain amastigotes by light microscopy of Giemsa-stained slides. The infected macrophage culture was scraped from the culture plates, centrifuged at 380g for 10 minutes and resuspended in heat-inactivated rabbit blood for sand fly infections at the concentration of 3 x 10^6 macrophages/ml.

Quantitative PCR

RNA was purified from individual midguts stored in –20 °C using a High Pure RNA Tissue Kit (Roche) and cDNA synthesized using SuperScript III Reverse Transcriptase (Invitrogen) with random hexamer primers (Promega) following manufacturer's instructions. Quantification of putative trypsin transcripts was performed by real-time PCR with primers designed for PperTryp1 (5'-CCC AAT GGA CTA TGA CTA CGC-3' and 5'-CGA ACA TCG TCG AAT ACG ATA G-3'), PperTryp2 (5'-GGT GTT CTC GTT GGA GTG GT-3' and 5'-TGG CGT AAA CTC CAG GGT AG-3'), PperTryp3 (5'-TGA GGA TGT TGA GGA TGG AA-3' and 5'-CTC TTG GTT ATT CAG AGT GTT ACC C-3') and PperS7 ribosomal protein as a reference transcript (5'-ATC CCT ATG CCG AAG CAG A-3' and 5'-TCA AGC TCA CGT ACC AGA CG-3'). The amplification reaction was carried out using the iQ5 real-time PCR detection system (Bio-Rad) by using the SYBR Green detection method (iQ SYBR Green Supermix, Bio-Rad) in 15µl reaction volume containing 1µl cDNA template and 0.5µl primer set (5uM each). The running conditions were as follows: 3 min at 95 °C followed by 40 repetitive cycles: 10 s at 95 °C, 10 s at 55 °C, and 10 s at 72 °C. Reactions were run in duplicates and data were analysed using the $2^{-\Delta\Delta CT}$ (Livak) method. In a preliminary experiment we have established that the target genes (trypsins) and the reference gene (PperS7 ribosomal protein) have similar (about 5% variance) amplification efficiencies which are nearly 100%. Data are presented as relative transcript levels using the S7 ribosomal protein gene as an internal control. Statistical analysis was done by Mann-Whitney U Test using (Statistica 6.1 StatSoft software).

Authors' contributions

AD participated in the study design, sand fly rearing, dissections, cDNA library construction and annotation, sequence alignment, phylogenetic analysis, quantitative PCR and drafting the manuscript. AJF and KDB sequenced all cDNA amplification products selected from the library. JV participated in the quantitative PCR experiment. PV and JGV participated in the study design and coordination, and revised the manuscript. RCJ conceived the study, participated in its design and coordination and revised the manuscript.

Acknowledgements and Funding

This work was supported by the Intramural Research Program of the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, the Wellcome trust project 078937, Czech Science Foundation (projects 206/09/0777 and 206/09/H026), Czech Ministry of Education (projects MSM0021620828 and LC06009) and the Boehringer Ingelheim Fonds. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade

names, commercial products, or organization imply endorsement by the government of the United States of America.

Reference List

1. **Killick-Kendrick R: The biology and control of phlebotomine sand flies. *Clinics In Dermatology* 1999, 17: 279-289.**

2. **Kamhawi S: Phlebotomine sand flies and Leishmania parasites: friends or foes? *Trends in Parasitology* 2006, 22: 439-445.**

3. **Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA *et al.*: Comparative genomic analysis of three Leishmania species that cause diverse human disease. *Nat Genet* 2007, 39: 839-847.**

4. **Oliveira F, Jochim RC, Valenzuela JG, Kamhawi S: Sand flies, Leishmania, and transcriptome-borne solutions. *Parasitol Int* 2009, 58: 1-5.**

5. **Ramalho-Ortigao M, Jochim RC, Anderson JM, Lawyer PG, Pham VM, Kamhawi S *et al.*: Exploring the midgut transcriptome of Phlebotomus papatasi: comparative analysis of expression profiles of sugar-fed, blood-fed and Leishmania major-infected sandflies. *Bmc Genomics* 2007, 8.**

6. **Jochim RC, Teixeira CR, Laughinghouse A, Mu JB, Oliveira F, Gomes RB *et al.*: The midgut transcriptome of Lutzomyia longipalpis: comparative analysis of cDNA libraries from sugar-fed, blood-fed, post-digested and Leishmania infantum chagasi-infected sand flies. *Bmc Genomics* 2008, 9.**

7. **Pitaluga AN, Beteille V, Lobo AR, Ortigao-Farias JR, Davila AM, Souza AA *et al.*: EST sequencing of blood-fed and Leishmania-infected midgut of Lutzomyia longipalpis, the principal visceral leishmaniasis vector in the Americas. *Mol Genet Genomics* 2009, 282: 307-317.**

8. **Ramalho-Ortigao JM, Kamhawi S, Joshi MB, Reynoso D, Lawyer PG, Dwyer DM *et al.*: Characterization of a blood activated chitinolytic system in the midgut of the sand fly vectors Lutzomyia longipalpis and Phlebotomus papatasi. *Insect Molecular Biology* 2005, 14: 703-712.**

9. **Kamhawi S, Ramalho-Ortigao M, Pham VM, Kumar S, Lawyer PG, Turco SJ *et al.*: A role for insect galectins in parasite survival. *Cell* 2004, 119: 329-341.**

10. **Sant'Anna MR, az-Albiter H, Mubaraki M, Dillon RJ, Bates PA: Inhibition of trypsin expression in Lutzomyia longipalpis using RNAi enhances the survival of Leishmania. *Parasit Vectors* 2009, 2: 62.**

11. **Pimenta PFP, Saraiva EMB, Rowton E, Modi GB, Garraway LA, Beverley SM *et al.*: Evidence that the vectorial competence of phlebotomine sand flies for different species of Leishmania is controlled by structural polymorphisms in the surface lipophosphoglycan. *Proc Natl Acad Sci USA* 1994, 91: 9155-9159.**

12. **Myskova J, Svobodova M, Beverley SM, Volf P: A lipophosphoglycan-independent development of Leishmania in permissive sand flies. *Microbes And Infection* 2007, 9: 317-324.**

13. **Svarovska A, Ant TH, Seblova V, Jecna L, Beverley SM, Volf P: Leishmania major glycosylation mutants require phosphoglycans (lpg2-) but not lipophosphoglycan (lpg1-) for survival in permissive sand fly vectors.** ***PLoS Negl Trop Dis*** **2010, 4: e580.**

14. **Guo Y, Ribeiro JM, Anderson JM, Bour S: dCAS: a desktop application for cDNA sequence annotation.** ***Bioinformatics*** **2009, 25: 1195-1196.**

15. **Ramalho-Ortigao JM, Kamhawi S, Rowton ED, Ribeiro JMC, Valenzuela JG: Cloning and characterization of trypsin- and chymotrypsin-like proteases from the midgut of the sand fly vector Phlebotomus papatasi.** ***Insect Biochemistry And Molecular Biology*** **2003, 33: 163-171.**

16. **Telleria EL, Pitaluga AN, Ortigao-Farias JR, de Araujo APO, Ramalho-Ortigao JM, Traub-Cseko YM: Constitutive and blood meal-induced trypsin genes in Lutzomyia longipalpis.** ***Archives of Insect Biochemistry and Physiology*** **2007, 66: 53-63.**

17. **Dillon RJ, Lane RP: Bloodmeal Digestion in the Midgut of Phlebotomus papatasi and Phlebotomus langeroni.** ***Med Vet Entomol*** **1993, 7: 225-232.**

18. **Schlein Y, Romano H: Leishmania major and L.donovani: effects on proteolytic enzymes of Phlebotomus papatasi (Diptera: Psychodidae).** ***Exp Parasitol*** **1986, 62: 376-380.**

19. **Dillon RJ, Lane RP: Influence of Leishmania Infection on Blood-Meal Digestion in the Sandflies Phlebotomus-papatasi and P.langeroni.** ***Parasitol Res*** **1993, 79: 492-496.**

20. **GRAF R, Briegel H: THE SYNTHETIC PATHWAY OF TRYPSIN IN THE MOSQUITO AEDES-AEGYPTI L (DIPTERA, CULICIDAE) AND INVITRO STIMULATION IN ISOLATED MIDGUTS.** ***Insect Biochemistry*** **1989, 19: 129-137.**

21. **Titani K, Ericsson LH, Walsh KA, Neurath H: Amino-acid sequence of bovine carboxypeptidase B.** ***Proc Natl Acad Sci U S A*** **1975, 72: 1666-1670.**

22. **Lavazec C, Boudin C, Lacroix R, Bonnet S, Diop A, Thiberge S** ***et al.*****: Carboxypeptidases B of Anopheles gambiae as targets for a Plasmodium falciparum transmission-blocking vaccine.** ***Infection And Immunity*** **2007, 75: 1635-1642.**

23. **Dinglasan RR, Kalume DE, Kanzok SM, Ghosh AK, Muratova O, Pandey A** ***et al.*****: Disruption of Plasmodium falciparum development by antibodies against a conserved mosquito midgut antigen.** ***Proceedings Of The National Academy Of Sciences Of The United States Of America*** **2007, 104: 13461-13466.**

24. **Abdullah MA, Valaitis AP, Dean DH: Identification of a Bacillus thuringiensis Cry11Ba toxin-binding aminopeptidase from the mosquito, Anopheles quadrimaculatus.** ***BMC Biochem*** **2006, 7:16.: 16.**

25. Pomes A, Melen E, Vailes LD, Retief JD, Arruda LK, Chapman MD: Novel allergen structures with tandem amino acid repeats derived from German and American cockroach. *J Biol Chem* 1998, 273: 30801-30807.

26. Wittstock U, Agerbirk N, Stauber EJ, Olsen CE, Hippler M, Mitchell-Olds T *et al.*: Successful herbivore attack due to metabolic diversion of a plant chemical defense. *Proc Natl Acad Sci U S A* 2004, 101: 4859-4864.

27. Fischer HM, Wheat CW, Heckel DG, Vogel H: Evolutionary origins of a novel host plant detoxification gene in butterflies. *Mol Biol Evol* 2008, 25: 809-820.

28. Abraham EG, Islam S, Srinivasan P, Ghosh AK, Valenzuela JG, Ribeiro JM *et al.*: Analysis of the Plasmodium and Anopheles transcriptional repertoire during ookinete development and midgut invasion. *J Biol Chem* 2004, 279: 5573-5580.

29. Shao L, Devenport M, Fujioka H, Ghosh A, Jacobs-Lorena M: Identification and characterization of a novel peritrophic matrix protein, Ae-Aper50, and the microvillar membrane protein, AEG12, from the mosquito, Aedes aegypti. *Insect Biochemistry And Molecular Biology* 2005, 35: 947-959.

30. Zaidman-Remy A, Herve M, Poidevin M, Pili-Floury S, Kim MS, Blanot D *et al.*: The Drosophila amidase PGRP-LB modulates the immune response to bacterial infection. *Immunity* 2006, 24: 463-473.

31. Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C *et al.*: Immunity-related genes and gene families in Anopheles gambiae. *Science* 2002, 298: 159-165.

32. Meister S, Agianian B, Turlure F, Relogio A, Morlais I, Kafatos FC *et al.*: Anopheles gambiae PGRPLC-mediated defense against bacteria modulates infections with malaria parasites. *PLoS Pathog* 2009, 5: e1000542.

33. Warr E, Das S, Dong Y, Dimopoulos G: The Gram-negative bacteria-binding protein gene family: its role in the innate immune system of anopheles gambiae and in anti-Plasmodium defence. *Insect Mol Biol* 2008, 17: 39-51.

34. Kumar S, Molina-Cruz A, Gupta L, Rodrigues J, Barillas-Mury C: A peroxidase/dual oxidase system modulates midgut epithelial immunity in Anopheles gambiae. *Science* 2010, 327: 1644-1648.

35. Boulanger N, Lowenberger C, Volf P, Ursic R, Sigutova L, Sabatier L *et al.*: Characterization of a defensin from the sand fly Phlebotomus duboscqi induced by challenge with bacteria or the protozoan parasite Leishmania major. *Infection And Immunity* 2004, 72: 7140-7146.

36. Kumar S, Christophides GK, Cantera R, Charles B, Han YS, Meister S *et al.*: The role of reactive oxygen species on Plasmodium melanotic

encapsulation in Anopheles gambiae. ***Proc Natl Acad Sci U S A*** **2003, 100: 14139-14144.**

37. Zdobnov EM, von MC, Letunic I, Torrents D, Suyama M, Copley RR *et al*.: Comparative genome and proteome analysis of Anopheles gambiae and Drosophila melanogaster. ***Science*** **2002, 298: 149-159.**

38. Kanzok SM, Fechner A, Bauer H, Ulschmid JK, Muller HM, Botella-Munoz J *et al*.: Substitution of the thioredoxin system for glutathione reductase in Drosophila melanogaster. ***Science*** **2001, 291: 643-646.**

39. Dillon RJ, Ivens AC, Churcher C, Holroyd N, Quail MA, Rogers ME *et al*.: Analysis of ESTs from Lutzomyia longipalpis sand flies and their contribution toward understanding the insect-parasite relationship. ***Genomics*** **2006, 88: 831-840.**

40. Jaramillo-Gutierrez G, Rodrigues J, Ndikuyeze G, Povelones M, Molina-Cruz A, Barillas-Mury C: Mosquito immune responses and compatibility between Plasmodium parasites and anopheline mosquitoes. ***BMC Microbiol*** **2009, 9:154.: 154.**

41. NARASIMHAN S, Sukumaran B, Bozdogan U, Thomas V, Liang X, Deponte K *et al*.: A tick antioxidant facilitates the Lyme disease agent's successful migration from the mammalian host to the arthropod vector. ***Cell Host Microbe*** **2007, 2: 7-18.**

42. Molina-Cruz A, DeJong RJ, Charles B, Gupta L, Kumar S, Jaramillo-Gutierrez G *et al*.: Reactive oxygen species modulate Anopheles gambiae immunity against bacteria and Plasmodium. ***J Biol Chem*** **2008, 283: 3217-3223.**

43. Sant'Anna MR, Alexander B, Bates PA, Dillon RJ: Gene silencing in phlebotomine sand flies: Xanthine dehydrogenase knock down by dsRNA microinjections. ***Insect Biochem Mol Biol*** **2008, 38: 652-660.**

44. Dunkov B, Georgieva T: Insect iron binding proteins: Insights from the genomes. ***Insect Biochemistry And Molecular Biology*** **2006, 36: 300-309.**

45. Ramalho-Ortigao JM, Traub-Cseko YM: Molecular characterization of Llchit1, a midgut chitinase cDNA from the leishmaniasis vector Lutzomyia longipalpis. ***Insect Biochemistry And Molecular Biology*** **2003, 33: 279-287.**

46. Sadlova J, Volf P: Peritrophic matrix of Phlebotomus duboscqi and its kinetics during Leishmania major development. ***Cell Tissue Res*** **2009, 337: 313-325.**

47. Dinglasan RR, Devenport M, Florens L, Johnson JR, McHugh CA, Donnelly-Doman M *et al*.: The Anopheles gambiae adult midgut peritrophic matrix proteome. ***Insect Biochem Mol Biol*** **2009, 39: 125-134.**

48. Lehane MJ: Peritrophic matrix structure and function. ***Annu Rev Entomol*** **1997.**

49. Devenport M, Alvarenga PH, Shao L, Fujioka H, Bianconi ML, Oliveira PL *et al*.: Identification of the Aedes aegypti peritrophic matrix protein AeIMUCI as a heme-binding protein. *Biochemistry* 2006, 45: 9540-9549.

50. Anderson JM, Oliveira F, Kamhawi S, Mans BJ, Reynoso D, Seitz AE *et al*.: Comparative salivary gland transcriptomics of sandfly vectors of visceral leishmaniasis. *Bmc Genomics* 2006, 7: 52.

51. Volf P, Volfova V: Establishment and maintenance of sand fly colonies. *Journal Of Vector Ecology* 2011, 36.

52. Francischetti IM, Calvo E, Andersen JF, Pham VM, Favreau AJ, Barbian KD *et al*.: Insight into the Sialome of the Bed Bug, Cimex lectularius. *J Proteome Res* 2010, 9: 3820-3831.

53. Huang X, Madan A: CAP3: A DNA sequence assembly program. *Genome Res* 1999, 9: 868-877.

54. Ewing B, Hillier L, Wendl MC, Green P: Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 1998, 8: 175-185.

55. Ewing B, Green P: Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 1998, 8: 186-194.

56. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W *et al*.: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997, 25: 3389-3402.

57. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM *et al*.: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000, 25: 25-29.

58. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV *et al*.: The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 2003, 4:41. Epub;%2003 Sep 11.: 41.

59. Schultz J, Copley RR, Doerks T, Ponting CP, Bork P: SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res* 2000, 28: 231-234.

60. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL: The Pfam protein families database. *Nucleic Acids Res* 2000, 28: 263-266.

61. Bendtsen JD, Nielsen H, von HG, Brunak S: Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004, 340: 783-795.

62. Sonnhammer EL, von HG, Krogh A: A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* 1998, 6:175-82.: 175-182.

63. Julenius K, Molgaard A, Gupta R, Brunak S: Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 2005, 15: 153-164.

64. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H *et al*.: Clustal W and Clustal X version 2.0. *Bioinformatics* 2007, 23: 2947-2948.

65. Abascal F, Zardoya R, Posada D: ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* 2005, 21: 2104-2105.

66. Schmidt HA, Strimmer K, Vingron M, von HA: TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 2002, 18: 502-504.

67. Tamura K, Dudley J, Nei M, Kumar S: MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007, 24: 1596-1599.

Figure Legends

Figure 1. Distribution of clusters from the sugar fed and blood fed libraries in general functional classes. Significant match to the KOG database (E<10E-5) was used as a guideline for grouping the sequences into the functional classes.

Figure 2. Multiple sequence alignment of putative sand fly trypsins. Pper: *Phlebotomus perniciosus*, Pp: *Phlebotomus papatasi*, Lulo: *Lutzomyia longipalpis*. Predicted signal peptides are underlined, the putative activation cleavage site is indicated by (↓), conserved cysteines (C), catalytic H/D/S residues marked by (*) and substrate binding site marked by (#). Accession numbers: PperTryp1 (EZ933288), PperTryp2 (EZ933289), PperTryp3 (EZ933290), Lltryp1 (ABM26904), Lltryp2 (ABM26905), LuloTryp3 (ABV60308), LuloTryp4 (ABV60300), PpTryp1 (AAM96940), PpTryp2 (AAM96941), PpTryp3 (AAM96942), PpTryp4 (AAM96943).

Figure 3. Influence of blood feeding and *L. infantum* infection on the dynamics of *P. perniciosus* trypsins expression. (A) The graph shows PperTryp1, PperTryp2 and PperTryp3 expression as fold over the reference housekeeping gene (PpPerS7 ribosomal protein) before and after the blood feeding (6 hours, 24hours , 72hours and 10 days). Each column represents the mean of ten females. S, sugar fed sand flies; B, blood fed sand flies; I, blood fed and *L. infantum* infected sand flies. The statistically significant difference between the infected and uninfected sand flies is indicated by (*). (B) The graph shows significant difference of PperTryp3 expression in uninfected (B_24) and infected (I_24) sand flies 24 hours after blood feeding; Mann-Whitney U Test (U=20; Z=2.268), p=0.023.

Figure 4. Phylogenetic analysis of putative trypsins from *Anopheles stephensi* (As), *Ochlerotatus taeniorhynchus* (Ot), *Homo sapiens* (Hs), *Aedes aegypti* (Aa), *Anopheles gambiae* (Ag), *Phlebotomus perniciosus* (Pper), *Lutzomyia longipalpis* (Lulo), *Phlebotomus papatasi* (Pp) and *Drosophila melanogaster* (Dm). Accession numbers are given in parentheses and node support is indicated by the bootstrap values.

Figure 5. Phylogenetic analysis of putative chymotrypsin molecules from *Litopenaeus vannamei* (Lv), *Culex quinquefasciatus* (Cq), *Anopheles gambiae* (Ag), *Aedes aegypti* (Ae), *Phlebotomus perniciosus* (Pper), *Lutzomyia longipalpis* (Lulo) and *Phlebotomus papatasi* (Pp). Accession numbers are given in parentheses and node support indicated by the bootstrap value.

Figure 6. Multiple sequence alignment of putative *P. perniciosus* chymotrypsins (partial sequences). Conserved cysteines are indicated (C), catalytic H/D/S residues marked by (*) and a serine residue implicated in chymotrypsin substrate specificity marked by (#). Accession numbers: PperChym1 (EZ933296), PperChym2 (EZ933297), PperChym3 (EZ933298), PperChym4 (EZ933299), PperChym5 (EZ933300).

Figure 7. Phylogenetic analysis of putative carboxypeptidases from *Drosophila melanogaster* (Dm), *Aedes aegypti* (Aa), *Anopheles gambiae* (Ag), *Phlebotomus perniciosus* (Pper), *Lutzomyia longipalpis* (Lulo) and *Phlebotomus papatasi* (Pp). Accession numbers are given in parentheses and node support is indicated by the bootstrap values.

Figure 8. Sequence alignment of putative midgut carboxypeptidases. (A) Comparison of mature Carboxypeptidase A proteins of *Aedes aegypti* (*A. aegypti*), *Anopheles gambiae* (*A. gambiae*), *Phlebotomus perniciosus* (Pper), *Phlebotomus papatasi* (Pp) and *Lutzomyia longipalpis* (Lulo). N-terminal portion of the peptides are not shown due to PperCpepA 5` mRNA truncation. Conserved cysteines are indicated (C), metal binding residues are marked by (*) and catalytic residues are marked by (#). (B) Comparison of mature Carboxypeptidase B proteins of *P. perniciosus* (Pper), *P. papatasi* (Pp) and *L. longipalpis* (Lulo). Conserved cysteines are indicated (C), metal binding residues are marked by (*), catalytic residues are marked by (#) and a conserved aspartate in the binding pocket of carboxypeptidases B is indicated (D). Accession numbers: PperCpepA (EZ966131), *A. aegypti* (AAT36730), *A. gambiae* (AAB96576), LuloCpepA1 **(**ABV60310), LuloCpepA2 (ABV60311), PpCpepA (ABV44738), PperCpepB (EZ966132), PpCpepB **(**ABV44754), LuloCpepB (ABV60312).

Figure 9. Phylogenetic analysis and sequence alignment of (putative) astacins. (A) *Astacus astacus* (Asa), *Drosophila melanogaster* (Dm), *Aedes aegypti* (Ae), *Anopheles gambiae* (Ag), *Phlebotomus perniciosus* (Pper), *Lutzomyia longipalpis* (Lulo), *Phlebotomus papatasi* (Pp) and *Culex quinquefasciatus* (Cq). Accession numbers are given in parentheses and node support is indicated by the bootstrap values. (B) *Drosophila melanogaster* (D. melanogaster), *Aedes aegypti* (*A. aegypti*), *Anopheles gambiae* (*A. gambiae*), *Phlebotomus perniciosus* (Pper), *Lutzomyia longipalpis* (Lulo), *Phlebotomus papatasi* (Pp) and *Culex quinquefasciatus* (*C. quinque*). Predicted signal peptide is underlined, conserved cysteines are indicated (C), the metal binding residues marked by (*) and catalytic residues marked by (#). Accession numbers: PperAstacin2 (ABV44746), PperAstacin1 (EZ966133), LuloAstacin (ABV60299), PpAstacin (ABV44746), *D. melanogaster* (AAY55427), *A. gambiae* (XP_318553), *A. aegypti* (XP_001648914), *C. quinque* (XP 001844556).

Figure 10. Phylogenetic analysis and sequence alignment of putative microvillar proteins. (A) *Periplaneta americana* (Pa), *Phlebotomus perniciosus* (Pper), *Phlebotomus papatasi* (Pp), *Lutzomyia longipalpis* (Lulo), *Aedes aegypti* (Ae), *Culex quinquefasciatus* (Cq) and *Anopheles gambiae* (Ag). Accession numbers are given in parentheses. Node support is indicated by the bootstrap values. (B) *Phlebotomus perniciosus* (Pper). The predicted signal peptides are underlined. Accession numbers: PperMVP1 (EZ933291), PperMVP2 (EZ933292), PperMVP3 (EZ933293), PperMVP4 (EZ933294), PperMVP5 (EZ933295).

Figure 11. Phylogenetic analysis of superoxide dismutase molecules from *Ixodes scapularis* (Is), *Anopheles gambiae* (Ag), *Aedes aegypti* (Ae), *Phlebotomus perniciosus* (Pper), *Lutzomyia longipalpis* (Lulo), *Phlebotomus papatasi* (Pp) and *Culex quinquefasciatus* (Cq). Accession numbers are given in parentheses, the clades are labelled with the respective localization based on SignalP prediction (Int: intracellular, Ext: extracellular) and node support is indicated by the bootstrap values.

Figure 12. Phylogenetic analysis of predicted chitin-binding domains of putative peritrophins from *Anopheles gambiae* (Ag), *Phlebotomus perniciosus* (Pper), *Lutzomyia longipalpis* (Lulo) and *Phlebotomus papatasi* (Pp). Accession numbers are given in parentheses and bootstrap values indicate node support.

Table 1: Selected clusters of combined *P. perniciosus* midgut cDNA libraries: putative function and sequence abundance in the sugar fed (SF) and blood fed (BF) libaries.

Cluster	Clone	GenBank	Name	Putative function	SF	BF	Total
46	PPRGUS_P2_F06	EZ933288	PperTryp1	trypsin	513	20	533
16	PPRGUM_P3_G08	EZ933289	PperTryp2	trypsin	10	0	10
63	PPRGFL_P8_E08	EZ933290	PperTryp3	trypsin	0	31	31
81	PPRGFL_P1_E01	EZ933296	PperChym1	chymotrypsin	0	82	82
102	PPRGFL_P7_D06	EZ933297	PperChym2	chymotrypsin	2	11	13
1033	PPRGUM_P1_A02	EZ933298	PperChym3	chymotrypsin	12	0	12
816	PPRGUM_P8_C08	EZ933299	PperChym4	chymotrypsin	7	0	7
710	PPRGUM_P7_F02	EZ933300	PperChym5	chymotrypsin	1	1	2
539	PPRGUL_P4_F08	EZ966131	PperCpepA	carboxypeptidase A	2	1	3
217	PPRGFL_P5_D07	EZ966132	PperCpepB	carboxypeptidase B	1	5	6
126	PPRGUL_P2_A02	EZ966135	PperApeptN	aminopeptidase N	7	13	20
84	PPRGFM_P8_D04	EZ966133	PperAstacin1	astacin	4	8	12
967	PPRGUL_P6_A03	EZ966134	PperAstacin2	astacin	3	0	3
45	PPRGFL_P1_A01	EZ933291	PperMVP1	microvillar protein	0	681	681
40	PPRGFL_P3_G12	EZ933292	PperMVP2	microvillar protein	0	26	26
18	PPRGFL_P5_H01	EZ933293	PperMVP3	microvillar protein	18	6	24
139	PPRGFL_P5_G10	EZ933294	PperMVP4	microvillar protein	0	28	28
52	PPRGFL_P1_E11	EZ933295	PperMVP5	microvillar protein	0	35	35
274	PPRGFL_P7_G05	EZ617708	PperPGRPLB	peptidoglycan recognition protein LB	1	1	2
168	PPRGFL_P5_C02	EZ617707	PperPGRPLC	peptidoglycan recognition protein LC	0	1	1
301	PPRGFL_P8_G01	EZ617706	PperGNBP	gram-negative bacteria binding protein	0	2	2
163	PPRGUM_P7_D04	EZ617709	PperGST1	glutathione-S-transferase Sigma	9	3	12
463	PPRGFM_P5_G03	EZ617710	PperGST2	microsomal glutathione-S-transferase	0	2	2
1322	PPRGUS_P3_F05	EZ617711	PperGST3	glutathione-S-transferase Theta	1	0	1
729	PPRGUM_P1_B05	HM119220	PperPRX	peroxiredoxin	1	0	1
852	PPRGUS-P1_C06	EZ617712	PperCat	catalase	9	0	9
892	PPRGUL_P4_B12	EZ617713	PperSOD1	Cu/Zn superoxide dismutase	1	0	1
1166	PPRGUM_P4_E03	EZ617714	PperSOD2	Cu/Zn superoxide dismutase	1	0	1
373	PPRGFM_P2_E03	EZ617715	PperXDH	xanthine dehydrogenase	1	1	2
88	PPRGFM_P2_G07	EZ617716	PperFLC	ferritin light chain	5	6	11
332	PPRGUS-P1_A12	EZ617717	PperFHC	ferritin heavy chain	9	7	16
330	PPRGFS_P8_D12	HM119221	PperPer2	peritrophin	1	4	5
156	PPRGUM_P7_A09	HM119222	PperPer3	peritrophin	1	2	3
97	PPRGFL_P6_E01	EZ933302	PperPer1	peritrophin	0	94	94
124	PPRGFL_P2_E05	EZ933285	PperChit	chitinase	0	3	3
358	PPRGUM_P6_A03	EZ933286	cluster 358	PM formation/unknown	10	3	13
379	PPRGFM_P6_A06	EZ933287	cluster 379	PM formation/unknown	2	4	6
174	PPRGFL_P4_A04	HQ015444	PperGH13	glycoside hydrolyse	23	9	32
183	PPRGFM_P6_G03	HQ015443	PperGH31	glycoside hydrolyse	12	3	15
79	PPRGFL_P6_G10	HQ015441	cluster 79	lipid recognition/unknown	0	9	9
652	PPRGUS_P7_D02	HQ015442	PperSA	40S ribosomal protein SA	8	1	9
461	PPRGUM_P6_G08	EZ933301	PperS7	ribosomal protein S7	6	2	8

Table 2: Selected clusters of combined *P. perniciosus* midgut cDNA libraries: best match to the NCBI non-redundant protein database.

Cluster	GenBank	Name	Best match to nr protein database	Best match GenBank	NR E value
46	EZ933288	PperTryp1	trypsin 2 [Lutzomyia longipalpis]	ABM26905.1	3,0E-82
16	EZ933289	PperTryp2	trypsin 1 [Phlebotomus papatasi]	AAM96940.1	3,0E-76
63	EZ933290	PperTryp3	putative trypsin 3 [Lutzomyia longipalpis]	ABV60308.1	1,0E-92
81	EZ933296	PperChym1	putative chymotrypsin [Lutzomyia longipalpis]	ABV60294.1	1,0E-109
102	EZ933297	PperChym2	chymotrypsin-like protein [Phlebotomus papatasi]	ABV44728.1	4,0E-94
1033	EZ933298	PperChym3	putative chymotrypsin [Lutzomyia longipalpis]	ABV60294.1	8,0E-77
816	EZ933299	PperChym4	putative chymotrypsin [Lutzomyia longipalpis]	ABV60293.1	6,0E-58
710	EZ933300	PperChym5	serine protease1/2 [Culex quinquefasciatus]	XP_001845462.1	1,0E-59
539	EZ966131	PperCpepA	carboxypeptidase A [Aedes aegypti]	AAT36730.1	1,0E-116
217	EZ966132	PperCpepB	carboxypeptidase B-like protein [Phlebotomus papatasi]	ABV44754.1	1,0E-170
126	EZ966135	PperApeptN	aminopeptidase N [Aedes aegypti]	AAK73351.1	1,0E-44
84	EZ966133	PperAstacin1	astacin-like metalloprotease [Lutzomyia longipalpis]	ABV60299.1	1,0E-92
967	EZ966134	PperAstacin2	AGAP010758-PA [Anopheles gambiae]	XP_318553.4	2,0E-49
45	EZ933291	PperMVP1	microvillar-like protein 1 [Lutzomyia longipalpis]	ABV60289.1	2,0E-69
40	EZ933292	PperMVP2	microvilli-like protein 2 [Phlebotomus papatasi]	ABV44759.1	7,0E-69
18	EZ933293	PperMVP3	microvilli-like protein 3 [Phlebotomus papatasi]	ABV44760.1	6,0E-57
139	EZ933294	PperMVP4	microvilli-like protein [Phlebotomus papatasi]	ABV44761.1	2,0E-92
52	EZ933295	PperMVP5	microvillar-like protein [Lutzomyia longipalpis]	ABV60295.1	1,0E-67
274	EZ617708	PperPGRPLB	putative PGRP [Phlebotomus papatasi]	ABV60369.1	1,0E-104
168	EZ617707	PperPGRPLC	PGRP-lc isoform [Anopheles gambiae]	AGAP005203-PC	4,0E-62
301	EZ617706	PperGNBP	GNBP [Aedes aegypti]	XP_001664288	1,0E-77
163	EZ617709	PperGST1	GST-like protein [Phlebotomus papatasi]	ABV44736.1	3,0E-113
463	EZ617710	PperGST2	microsomal GST [Culex quinquefasciatus]	XP_001863047.1	8,0E-20
1322	EZ617711	PperGST3	GST theta [Aedes aegypti]	XP_001659667.1	8,0E-22
729	HM119220	PperPRX	peroxiredoxin-like [Phlebotomus papatasi]	ABV44727.1	6,0E-86
852	EZ617712	PperCat	putative catalase [Lutzomyia longipalpis]	ABV60342.1	0,0E+00
892	EZ617713	PperSOD1	putative Cu/Zn SOD [Lutzomyia longipalpis]	ABV60343.1	5,0E-89
1166	EZ617714	PperSOD2	superoxide dismutase [Culex quinquefasciatus]	XP_001866335	9,0E-64
373	EZ617715	PperXDH	XDH [Lutzomyia longipalpis]	CAP08999.1	4,0E-45
88	EZ617716	PperFLC	FLC-like [Phlebotomus papatasi]	ABV44741.1	1,0E-111
332	EZ617717	PperFHC	FHC-like [Phlebotomus papatasi]	ABV44737	7,0E-73
330	HM119221	PperPer2	putative peritrophin [Lutzomyia longipalpis]	ABV60320.1	2,0E-09
156	HM119222	PperPer3	peritrophin-like protein [Phlebotomus papatasi]	ABV44751.1	4,0E-59
97	EZ933302	PperPer1	peritrophin-like protein [Phlebotomus papatasi]	ABV44705.1	1,0E-102
124	EZ933285	PperChit	midgut chitinase [Phlebotomus papatasi]	AAV49322.1	4,0E-61
358	EZ933286	cluster 358	14.5 kDa salivary protein [Phlebotomus duboscqi]	ABI20163	3,0E-49
379	EZ933287	cluster 379	31.5 kDa midgut protein [Phlebotomus papatasi]	ABV44721.1	5,0E-90
174	HQ015444	PperGH13	alpha-amylase [Aedes aegypti]	XP_001649787.1	1,0E-170
183	HQ015443	PperGH31	GK14321 [Drosophila willistoni]	XP_002073831.1	0,0E+00
79	HQ015441	cluster 79	Niemann-Pick Type C2, putative [Aedes aegypti]	XP_001647805.1	2,0E-14
652	HQ015442	PperSA	40S ribosomal protein SA [Simulium nigrimanum]	ACZ28384.1	1,0E-68
461	EZ933301	PperS7	40S ribosomal protein S7-like protein [Phlebotomus papatasi]	ABV44745.1	2,0E-95

Table 3: ESTs overrepresented in the blood fed library (BF) in comparison to the sugar fed library (SF)

Cluster	Putative function	SF	BF	P value
45	microvillar protein (PperMVP1)	0	681	7,49E-185
97	peritrophin (PperPer1)	0	94	4,46E-23
81	chymotrypsin (PperChym1)	0	82	2,83E-20
52	microvillar protein (PperMVP5)	0	35	2,10E-09
63	trypsin (PperTryp3)	0	31	1,76E-08
139	microvillar protein (PperMVP4)	0	28	8,67E-08
40	microvillar protein (PperMVP2)	0	26	2,52E-07
79	unknown (lipid recognition)	0	9	2,47E-03
102	chymotrypsin (PperChym2)	2	11	1,14E-02

Table 4: ESTs overrepresented in the sugar fed library (SF) in comparison to the blood fed library (BF)

Cluster	Putative function	SF	BF	P value
46	trypsin (PperTryp1)	513	20	3,16E-115
249	unknown	25	3	3,60E-05
1033	chymotrypsin (PperChym3)	12	0	5,76E-04
16	trypsin (PperTryp2)	10	0	1,68E-03
852	catalase (PperCat)	10	0	2,88E-03
174	glycoside hydrolase (PperGH13)	23	9	1,46E-02
18	microvillar protein (PperMVP3)	18	6	1,56E-02
652	40S ribosomal protein SA	8	1	2,07E-02
183	glycoside hydrolase (PperGH31)	12	3	2,16E-02

Figure 1

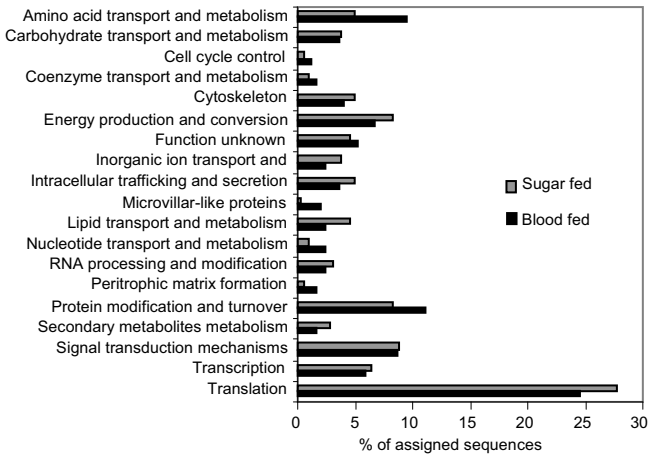

Amino acid transport and metabolism
Carbohydrate transport and metabolism
Cell cycle control
Coenzyme transport and metabolism
Cytoskeleton
Energy production and conversion
Function unknown
Inorganic ion transport and
Intracellular trafficking and secretion
Microvillar-like proteins
Lipid transport and metabolism
Nucleotide transport and metabolism
RNA processing and modification
Peritrophic matrix formation
Protein modification and turnover
Secondary metabolites metabolism
Signal transduction mechanisms
Transcription
Translation
Sugar fed
Blood fed
0
5
10
15
20
25
30
% of assigned sequences

Figure 2

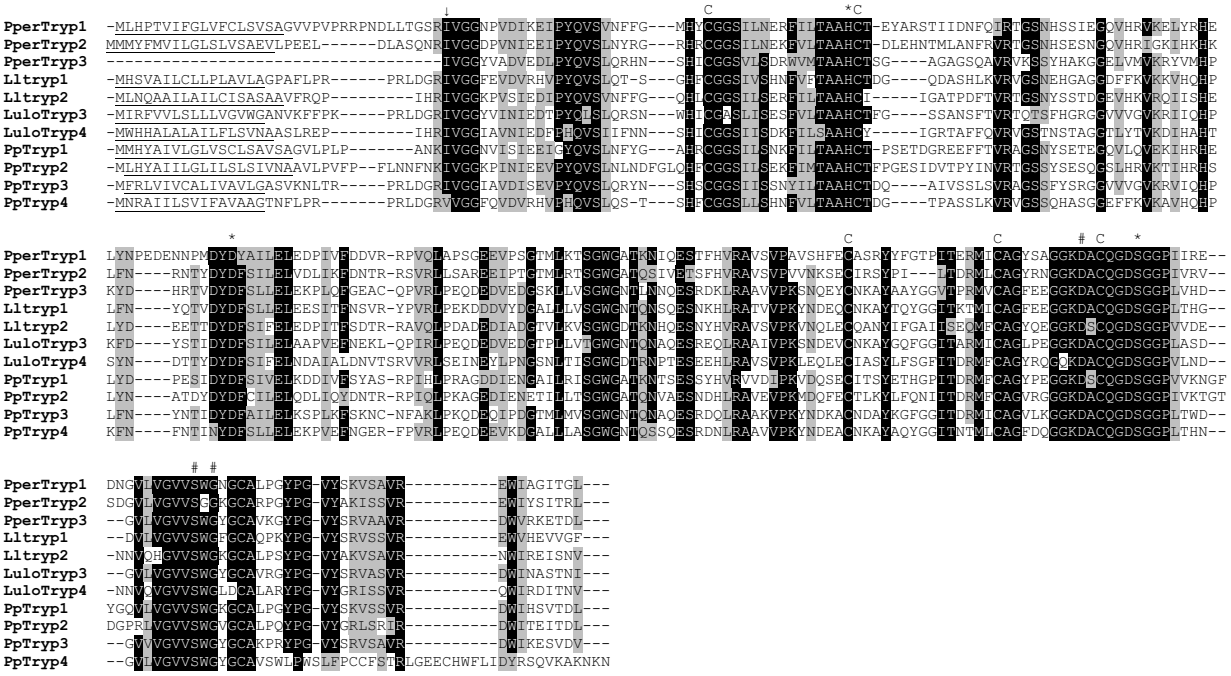

PperTryp1
PperTryp2
PperTryp3
Lltryp1
Lltryp2
LuloTryp3
LuloTryp4
PpTryp1
PpTryp2
PpTryp3
PpTryp4

Figure 3

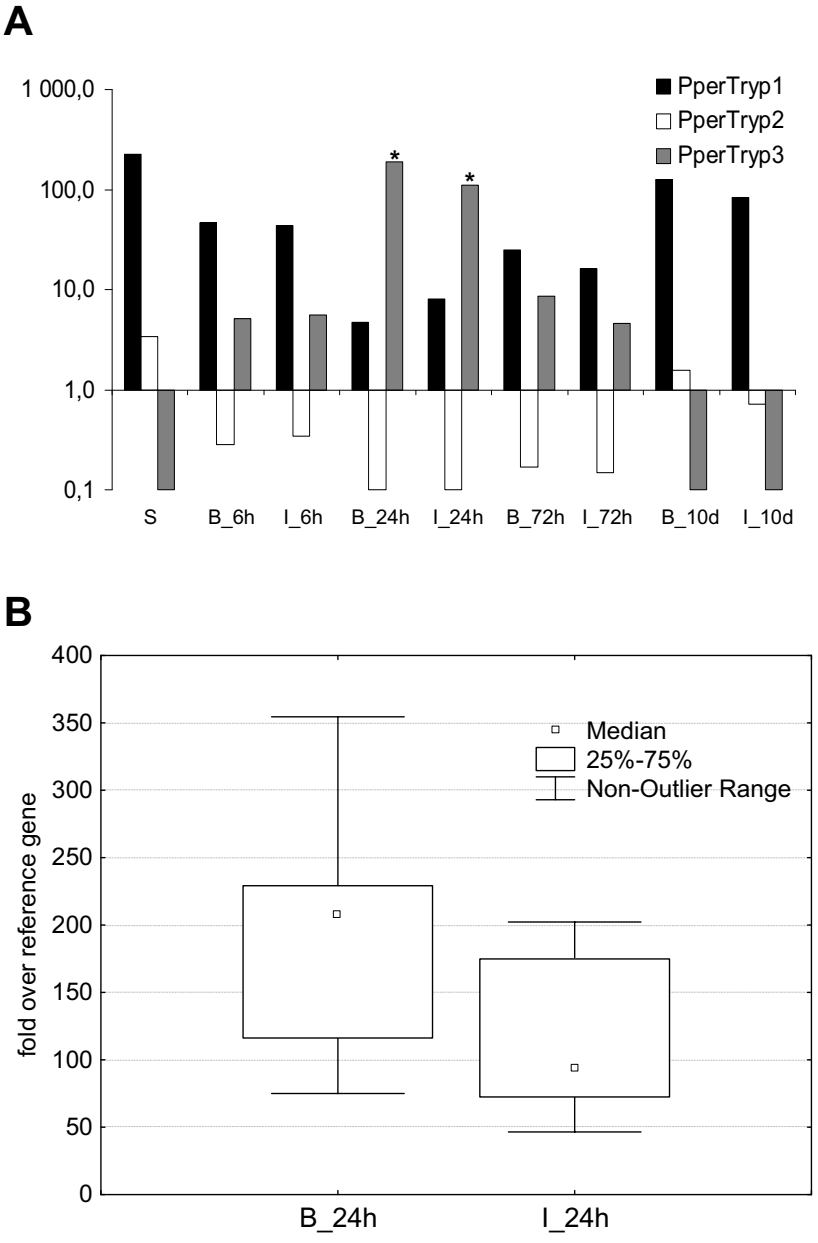

A
1 000,0
100,0
10,0
1,0
0,1
PperTryp1
PperTryp2
PperTryp3
*
*
S
B_6h
I_6h
B_24h
I_24h
B_72h
I_72h
B_10d
I_10d
B
400
350
300
250
200
150
100
50
0
fold over reference gene
Median
25%-75%
Non-Outlier Range
B_24h
I_24h

Figure 4

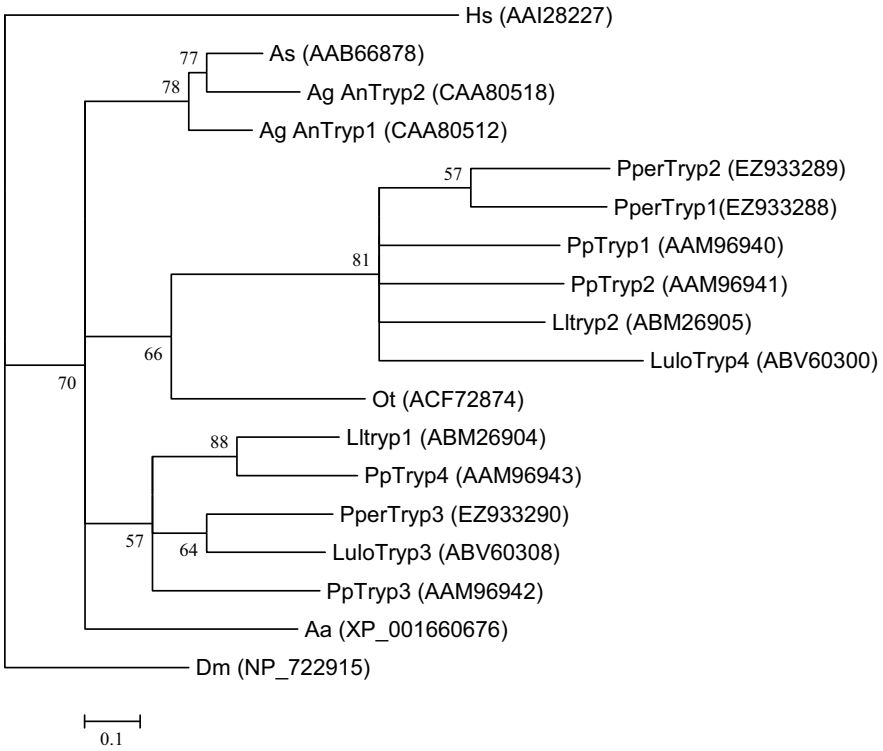

Hs (AAI28227)
77
As (AAB66878)
78
Ag AnTryp2 (CAA80518)
Ag AnTryp1 (CAA80512)
57
PperTryp2 (EZ933289)
PperTryp1(EZ933288)
81
PpTryp1 (AAM96940)
PpTryp2 (AAM96941)
Lltryp2 (ABM26905)
66
LuloTryp4 (ABV60300)
70
Ot (ACF72874)
88
Lltryp1 (ABM26904)
PpTryp4 (AAM96943)
PperTryp3 (EZ933290)
57
64
LuloTryp3 (ABV60308)
PpTryp3 (AAM96942)
Aa (XP_001660676)
Dm (NP_722915)
0.1

Figure 5

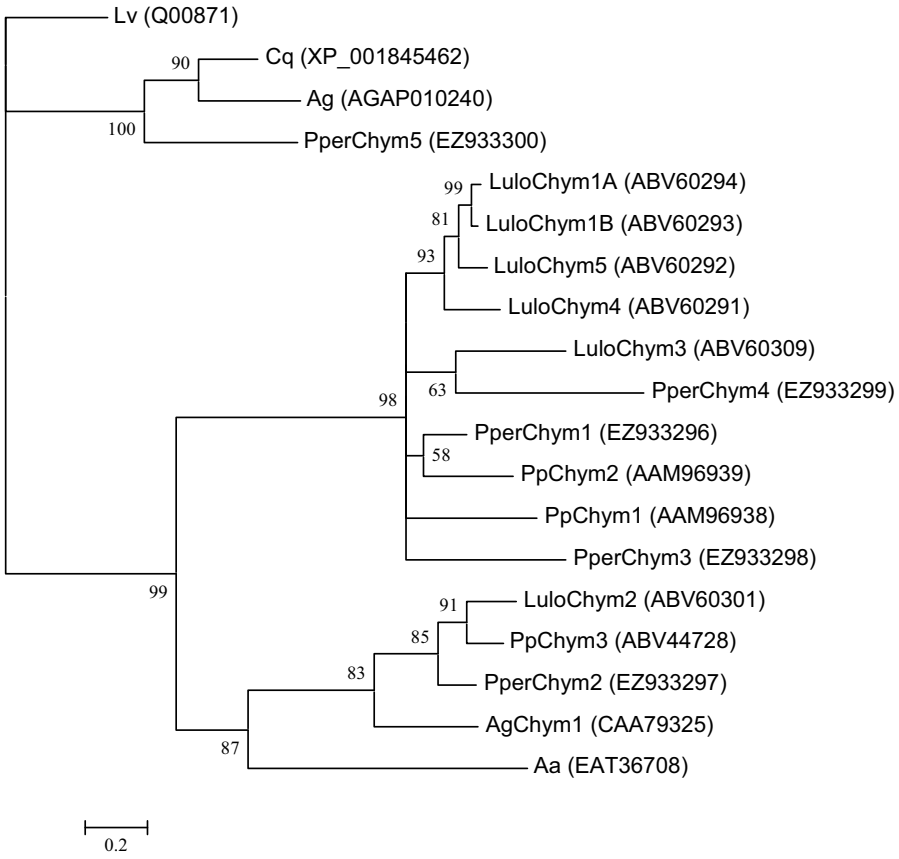

Lv (Q00871)
90
Cq (XP_001845462)
Ag (AGAP010240)
100
PperChym5 (EZ933300)
99
LuloChym1A (ABV60294)
81
LuloChym1B (ABV60293)
93
LuloChym5 (ABV60292)
LuloChym4 (ABV60291)
LuloChym3 (ABV60309)
63
PperChym4 (EZ933299)
98
PperChym1 (EZ933296)
58
PpChym2 (AAM96939)
PpChym1 (AAM96938)
PperChym3 (EZ933298)
99
91
LuloChym2 (ABV60301)
85
PpChym3 (ABV44728)
83
PperChym2 (EZ933297)
AgChym1 (CAA79325)
87
Aa (EAT36708)
0.2

Figure 6

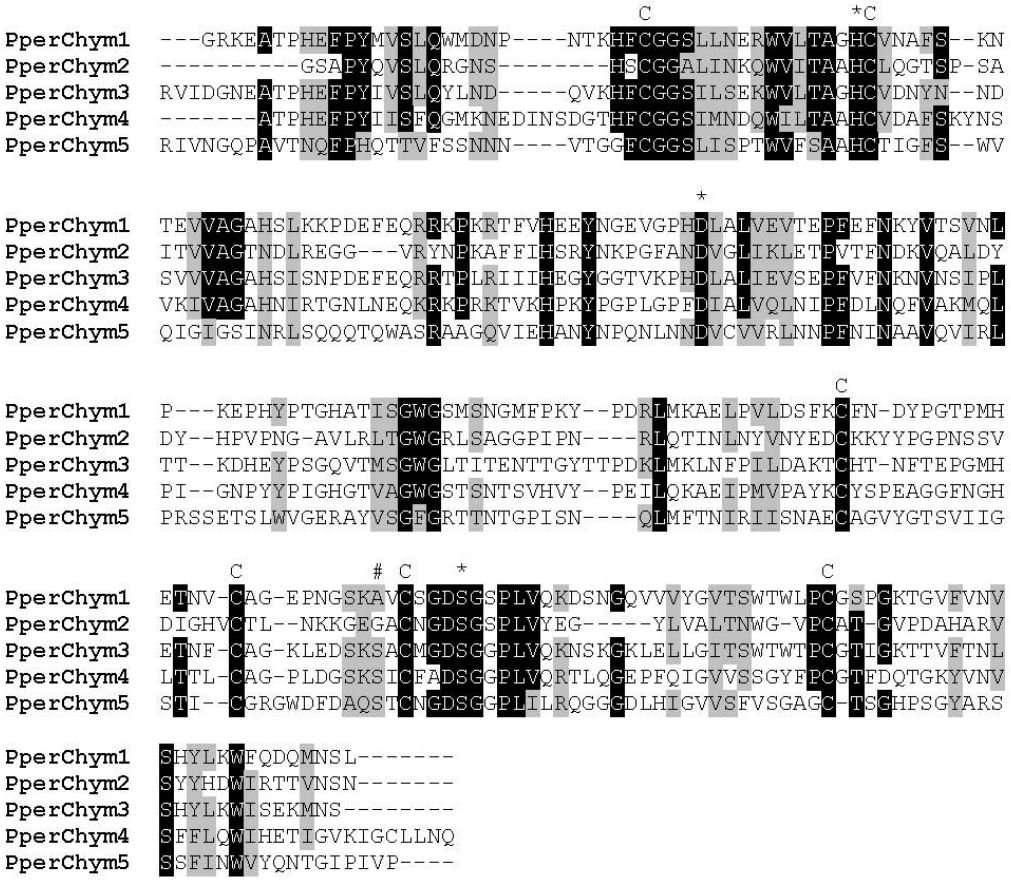


Figure 7

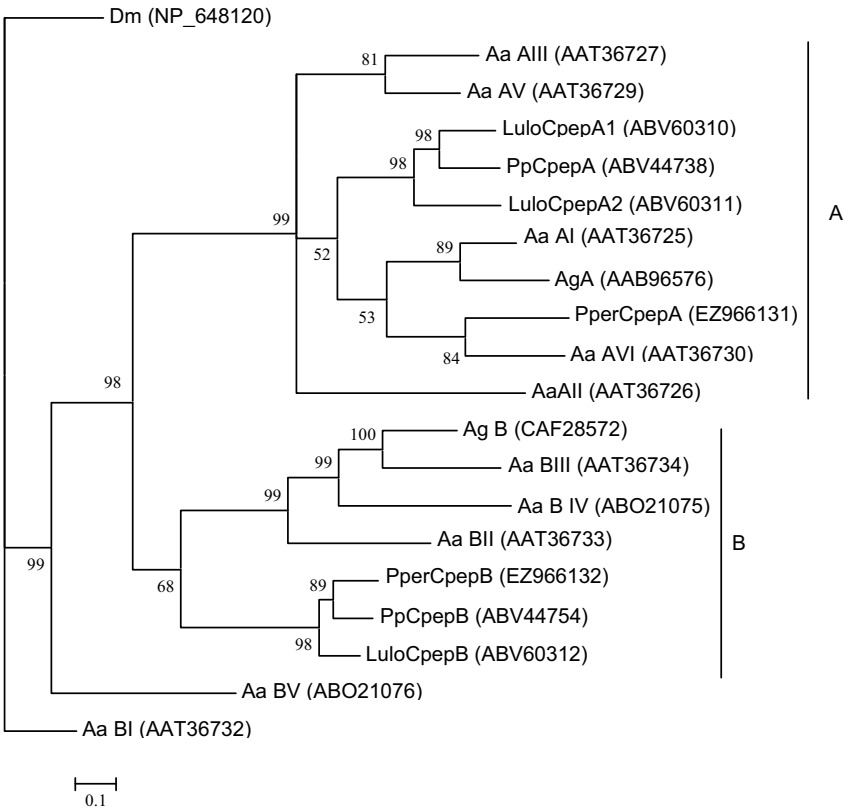

Dm (NP_648120)
Aa AIII (AAT36727)
Aa AV (AAT36729)
LuloCpepA1 (ABV60310)
PpCpepA (ABV44738)
LuloCpepA2 (ABV60311)
Aa AI (AAT36725)
AgA (AAB96576)
PperCpepA (EZ966131)
Aa AVI (AAT36730)
AaAII (AAT36726)
Ag B (CAF28572)
Aa BIII (AAT36734)
Aa B IV (ABO21075)
Aa BII (AAT36733)
PperCpepB (EZ966132)
PpCpepB (ABV44754)
LuloCpepB (ABV60312)
Aa BV (ABO21076)
Aa BI (AAT36732)
A
B
81
98
98
99
52
89
53
84
98
100
99
99
99
68
89
98
0.1

Figure 8

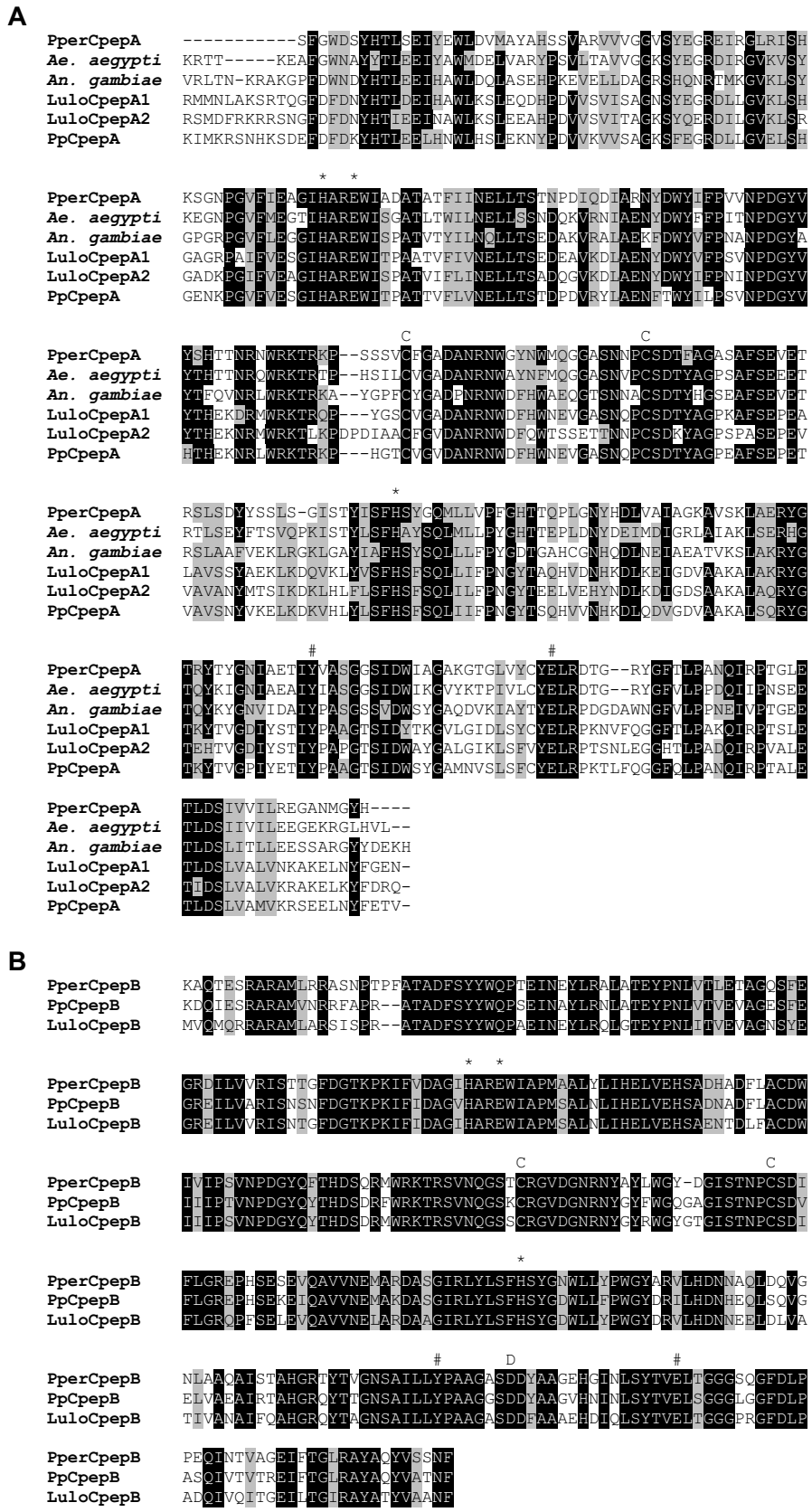
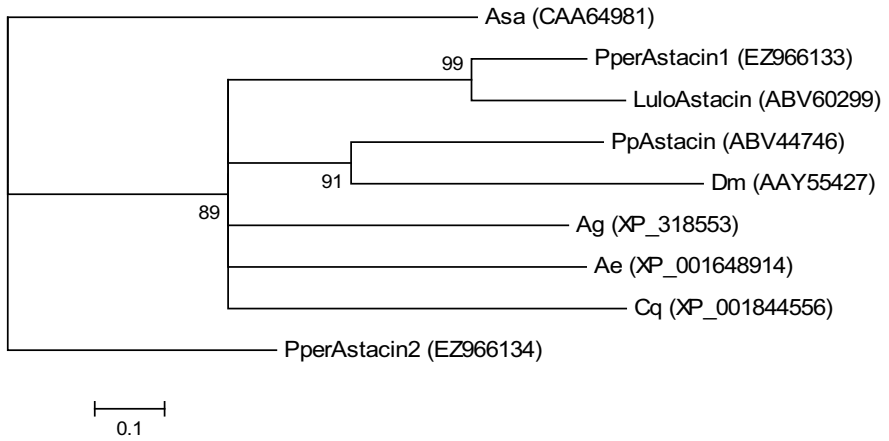

A
PperCpepA
Ae. aegypti
An. gambiae
LuloCpepA1
LuloCpepA2
PpCpepA
B
PperCpepB
PpCpepB
LuloCpepB

Figure 9

A


Asa (CAA64981)
99
PperAstacin1 (EZ966133)
LuloAstacin (ABV60299)
PpAstacin (ABV44746)
91
Dm (AAY55427)
89
Ag (XP_318553)
Ae (XP_001648914)
Cq (XP_001844556)
PperAstacin2 (EZ966134)
0.1

B

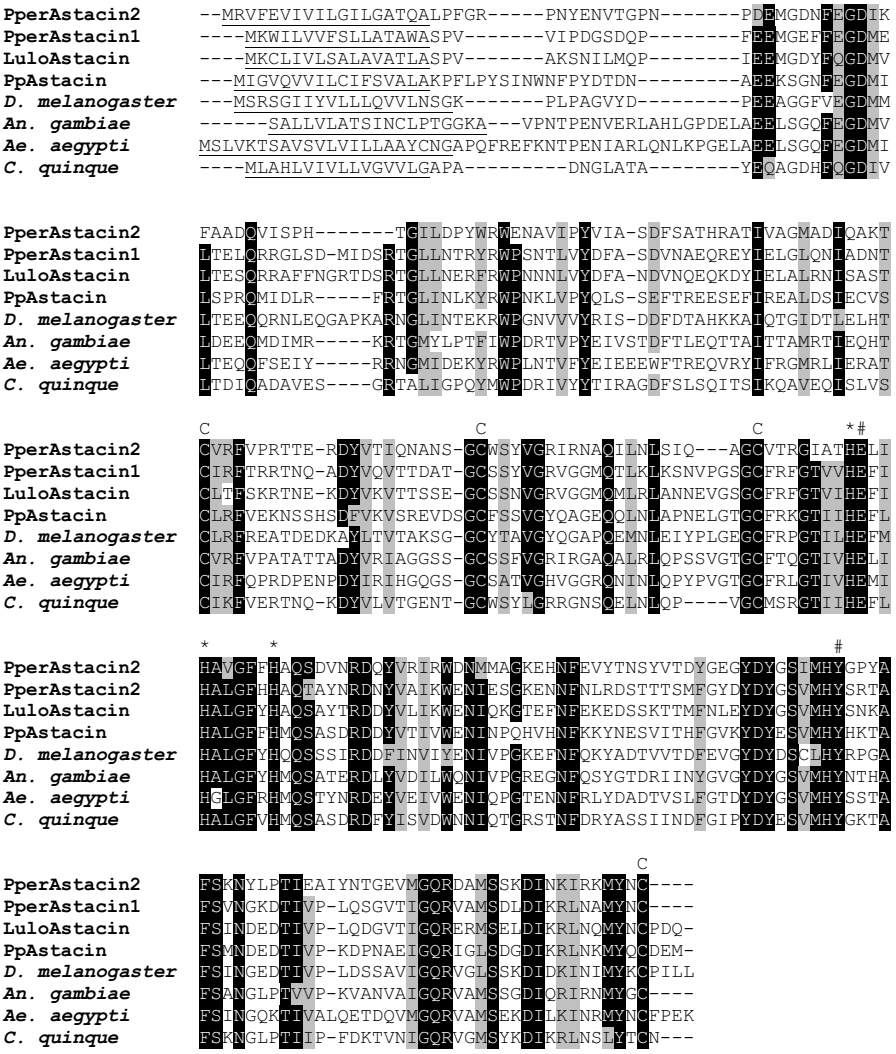
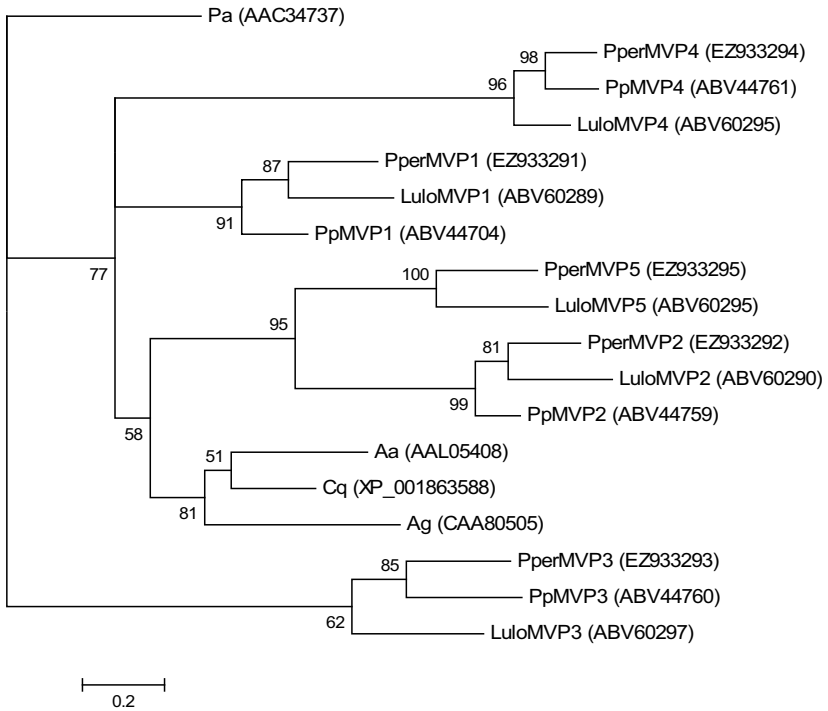

PperAstacin2 --MRVFEVIVILGILGATQALPFGR-----PNYENVTGPN-------PDEMGDNFEGDIK
PperAstacin1 ----MKWILVVFSLLATAWASPV-------VIPDGSDQP--------FEEMGEFFEGDME
LuloAstacin ----MKCLIVLSALAVATLASPV-------AKSNILMQP--------IEEMGDYFQGDMV
PpAstacin ---MIGVQVVILCIFSVALAKPFLPYSINWNFPYDTDN---------AEEKSGNFEGDMI
D. melanogaster ---MSRSGIIYVLLLQVVLNSGK-------PLPAGVYD---------PEEAGGFVEGDMM
An. gambiae ------SALLVLATSINCLPTGGKA---VPNTPENVERLAHLGPDELAEELSGQFEGDMV
Ae. aegypti MSLVKTSAVSVLVILLAAYCNGAPQFREFKNTPENIARLQNLKPGELAEELSGQFEGDMI
C. quinque ----MLAHLVIVLLVGVVLGAPA---------DNGLATA--------YEQAGDHFQGDIV
PperAstacin2 FAADQVISPH-------TGILDPYWRWENAVIPYVIA-SDFSATHRATIVAGMADIQAKT
PperAstacin1 LTELQRRGLSD-MIDSRTGLLNTRYRWPSNTLVYDFA-SDVNAEQREYIELGLQNIADNT
LuloAstacin LTESQRRAFFNGRTDSRTGLLNERFRWPNNNLVYDFA-NDVNQEQKDYIELALRNISAST
PpAstacin LSPRQMIDLR-----FRTGLINLKYRWPNKLVPYQLS-SEFTREESEFIREALDSIECVS
D. melanogaster LTEEQQRNLEQGAPKARNGLINTEKRWPGNVVVYRIS-DDFDTAHKKAIQTGIDTLELHT
An. gambiae LDEEQMDIMR-----KRTGMYLPTFIWPDRTVPYEIVSTDFTLEQTTAITTAMRTIEQHT
Ae. aegypti LTEQQFSEIY-----RRNGMIDEKYRWPLNTVFYEIEEEWFTREQVRYIFRGMRLIERAT
C. quinque LTDIQADAVES----GRTALIGPQYMWPDRIVYYTIRAGDFSLSQITSIKQAVEQISLVS
PperAstacin2 CVRFVPRTTE-RDYVTIQNANS-GCWSYVGRIRNAQILNLSIQ---AGCVTRGIATHELI
PperAstacin1 CIRFTRRTNQ-ADYVQVTTDAT-GCSSYVGRVGGMQTLKLKSNVPGSGCFRFGTVVHEFI
LuloAstacin CLTFSKRTNE-KDYVKVTTSSE-GCSSNVGRVGGMQMLRLANNEVGSGCFRFGTVIHEFI
PpAstacin CLRFVEKNSSHSDFVKVSREVDSGCFSSVGYQAGEQQLNLAPNELGTGCFRKGTIIHEFL
D. melanogaster CLRFREATDEDKAYLTVTAKSG-GCYTAVGYQGAPQEMNLEIYPLGEGCFRPGTILHEFM
An. gambiae CVRFVPATATTADYVRIAGGSS-GCSSFVGRIRGAQALRLQPSSVGTGCFTQGTIVHELI
Ae. aegypti CIRFQPRDPENPDYIRIHGQGS-GCSATVGHVGGRQNINLQPYPVGTGCFRLGTIVHEMI
C. quinque CIKFVERTNQ-KDYVLVTGENT-GCWSYIGRRGNSQELNLQP----VGCMSRGTIIHEFL
PperAstacin2 HAVGFFHAQSDVNRDQYVRIRWDNMMACKEHNFEVYTNSYVTDYGEGYDYGSIMHYGPYA
PperAstacin2 HALGFFHAQTAYNRDNYVAIKWENIESGKENNFNLRDSTTTSMFGYDYDYGSVMHYSRTA
LuloAstacin HALGFYHAQSAYTRDDYVLIKWENIQKGTEFNFEKEDSSKTTMFNLEYDYGSVMHYSNKA
PpAstacin HALGFFHMQSASDRDDYVTIVWENINPQHVHNFKKYNESVITHFGVKYDYESVMHYHKTA
D. melanogaster HALGFYHQQSSSIRDDFINVIYENIVPGKEFNFQKYADTVVTDFEVGYDYDSCLHYRPGA
An. gambiae HALGFYHMQSATERDDLYVDILWQNIVPGREGNFQSYGTDRIINYGVGYDYGSVMHYNTHA
Ae. aegypti HGLGFRHMQSTYNRDEYVEIVWENIQPGTENNFRLYDADTVSLFGTDYDYGSVMHYSSTA
C. quinque HALGFVHMQSASDRDFYISVDWNNIQTGRSTNFDRYASSIINDFGIPYDYESVMHYGKTA
PperAstacin2 FSKNYLPTIEAIYNTGEVMGQRDAMSSKDINKIRKMYNC----
PperAstacin1 FSVNGKDTIVP-LQSGVTIGQRVAMSDLDIKRLNAMYNC----
LuloAstacin FSINDEDTIVP-LQDGVTIGQRERMSELDIKRLNQMYNCPDQ-
PpAstacin FSMNDEDTIVP-KDPNAEIGQRIGLSDGDIKRLNKMYQCDEM-
D. melanogaster FSINGEDTIVP-LDSSAVIGQRVGLSSKDIDKINIMYKCPILL
An. gambiae FSANGLPTVVP-KVANVAIGQRVAMSSGDIQRIRNMYGC----
Ae. aegypti FSINGQKTIVALQETDQVMGQRVAMSEKDILKINRMYNCFPEK
C. quinque FSKNGLPTIIP-FDKTVNIGQRVGMSYKDIKRLNSIYTCN---

Figure 10

A



B

```
PperMVP1    MKFLVVFAIIAAVATSQA--------FTVEPRGLQEDLQEFQDDLPVEDMVKLAVDYLMN
PperMVP2    ------------------------------------EDFDDFVNLVDLVQVKRVALRYYAS
PperMVP3    ------------------------------------EDLQDFANLIPTTDIIGFVLTYAIT
PperMVP4    -LIAIFCLVCSVVAQSGLEEPSTFNLELFLTPELNKDFADFVALLPMDEVLQVAENHYEN
PperMVP5    MKFAVACILLTFCLSALA----------VTPRDLQQDFEEFNALIPSKKIQDVVTKYYLF

PperMVP1    DKEFQAVVAYFQS-AEEAKLWEGVFKNKEVKEVVAFLQNAGLDVVGLLNEVANFLNLEPV
PperMVP2    DKETKQFVKYLKG-DTFGAVWDQVFTNEHAFNFLKDLQNSGVNVESVVNQVAKFLKHEQV
PperMVP3    DPEVRNLYQYVSRSQEFRELYIRAMDSTAVHDLVSILESYNLPVVAAINQVAFLLALEDY
PperMVP4    NSGLNKTLNYLKT-NKEAKHWDNLFSLSEVNKFLDFVNNTGLNVFGVLNEFAEYFELSPV
PperMVP5    DGETRNFVKYLKG-AEERKVWDQVFTYPDVKDLLHYLKDKNADATGLLNSLADLLGLEHV

PperMVP1    KP----TN-----ARRGTGLEGFLNDLLAMLPKDKLVALFKEKLQTSPEFKAFYEKLTGA
PperMVP2    DFNSLLS------RKPGKGLKELVNEVLQHVPLDKMKKLSDEKMESSPEFKELYETVAGF
PperMVP3    VPRQAADFS--EVKAPTKGVNGLVDSVLRLLPREELIKLYIQKLATSPEFRQLNKNLASV
PperMVP4    GDDFEDEDLPEDRFEYTWGFNALVNDVVDLLPKDDLKALFDQKVAEGEDFANFVENYSTE
PperMVP5    TPDLMLGN-----AVMLRSLKGLFNEIVALLPLDKLQALLEDKLANSEDFKELYKRIASY

PperMVP1    EFKKLTDSMFNSKEVQGYLKELKNHKVDLLKFLQVFKEEFGFP---
PperMVP2    DYEAMKDFVANSEEMLDLFETVRSYGLDLDGFFHQVEEEFGWN---
PperMVP3    ETAKAVSRVLFNQEVRRAYSEVEKRGVNLVQIAKQVVVYELDL---
PperMVP4    EFRKLLKKLEMSPVAQKLFKRFRKHGLDVYKLVNLGFALFGLN---
PperMVP5    DYCALEQFTANSAEVQDFLNRLNGYGVSLDGLFNDVKEFFGWNKCN
```

Figure 11

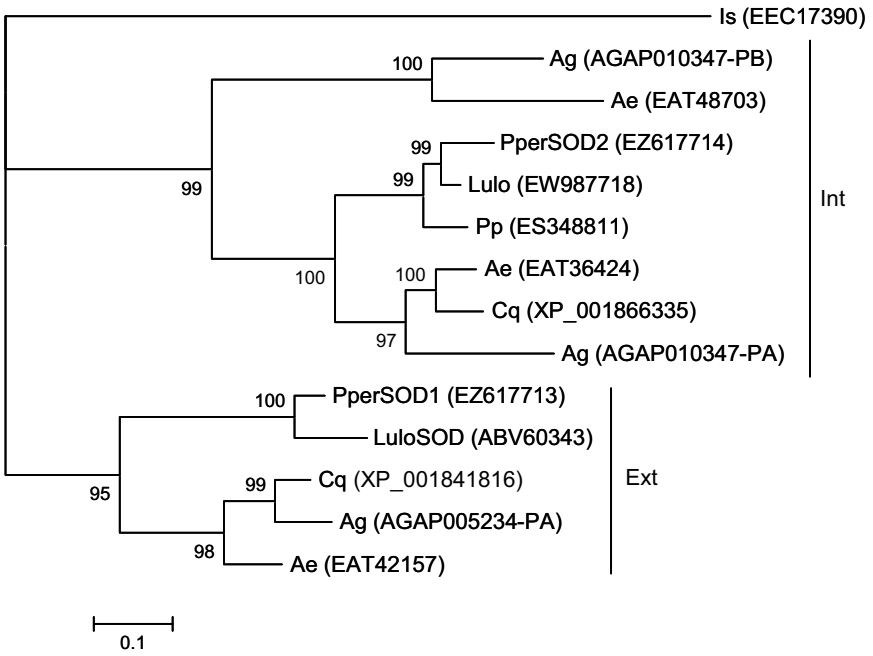
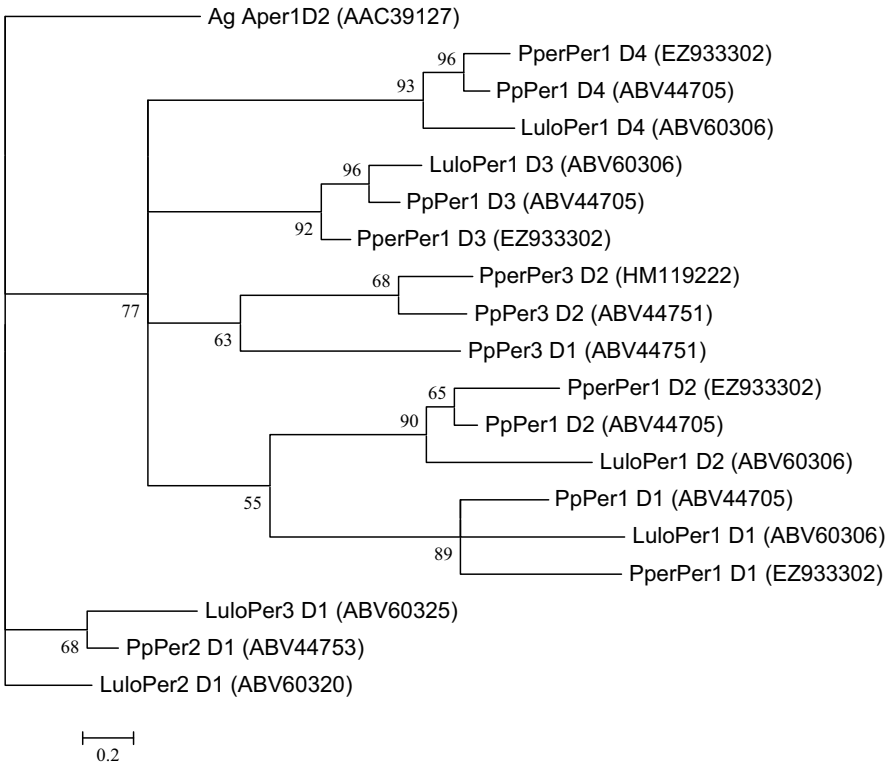

Is (EEC17390)
100
Ag (AGAP010347-PB)
Ae (EAT48703)
99
PperSOD2 (EZ617714)
99
Lulo (EW987718)
99
Pp (ES348811)
Int
100
100
Ae (EAT36424)
Cq (XP_001866335)
97
Ag (AGAP010347-PA)
100
PperSOD1 (EZ617713)
LuloSOD (ABV60343)
95
99
Cq (XP_001841816)
Ext
Ag (AGAP005234-PA)
98
Ae (EAT42157)
0.1

Figure 12


Ag Aper1D2 (AAC39127)
96
PperPer1 D4 (EZ933302)
93
PpPer1 D4 (ABV44705)
LuloPer1 D4 (ABV60306)
96
LuloPer1 D3 (ABV60306)
PpPer1 D3 (ABV44705)
92
PperPer1 D3 (EZ933302)
68
PperPer3 D2 (HM119222)
77
PpPer3 D2 (ABV44751)
63
PpPer3 D1 (ABV44751)
65
PperPer1 D2 (EZ933302)
90
PpPer1 D2 (ABV44705)
LuloPer1 D2 (ABV60306)
55
PpPer1 D1 (ABV44705)
LuloPer1 D1 (ABV60306)
89
PperPer1 D1 (EZ933302)
LuloPer3 D1 (ABV60325)
68
PpPer2 D1 (ABV44753)
LuloPer2 D1 (ABV60320)
0.2